RSC Medicinal Chemistry

ROYAL SOCIETY OF CHEMISTRY

RESEARCH ARTICLE

Cite this: *RSC Med. Chem.*, 2022, **13**, 1575

Revisiting a challenging p53 binding site: a diversity-optimized HEFLib reveals diverse binding modes in T-p53C-Y220C

Jason Stahlecker, [a] Theresa Klett, [a] Martin Schwer,[a] Simon Jaag,[b] Marcel Dammann,[a] Larissa N. Ernst,[a] Michael B. Braun, [c] Markus O. Zimmermann,[a] Markus Kramer,[d] Michael Lämmerhofer,[b] Thilo Stehle,[c] Murray Coles [e] and Frank M. Boeckler *[af]

Received 24th July 2022,
Accepted 5th September 2022

DOI: 10.1039/d2md00246a

rsc.li/medchem

The cellular tumor antigen p53 is a key component in cell cycle control. The mutation Y220C heavily destabilizes the protein thermally but yields a druggable crevice. We have screened the diversity-optimized halogen-enriched fragment library against T-p53C-Y220C with STD-NMR and DSF to identify hits, which we validated by $^{1}H,^{15}N$-HSQC NMR. We could identify four hits binding in the Y220C cleft, one hit binding covalently and four hits binding to an uncharacterized binding site. Compound 1151 could be crystallized showing a flip of C220 and thus opening subsite 3. Additionally, 4482 was identified to alkylate cysteines. Data shows that the diversity-optimized HEFLib leads to multiple diverse hits. The identified scaffolds can be used to further optimize interactions with T-p53C-Y220C and increase thermal stability.

1 Introduction

The cellular tumor antigen p53 plays a crucial role in cell cycle regulation.[1] Upregulation can lead to *e.g.* cell cycle arrest and apoptosis.[2,3] p53 acts as a tumor suppressor and is most frequently inactivated in cancer either through direct mutation of p53 or through perturbation of its associated pathways. As a consequence, reactivation of p53 function in tumors is perceived as a prime target for therapeutic intervention.[4–8]

The low thermal stability makes handling and thus research of p53 and/or the DNA binding domain (p53C, residues 94–312) challenging. A quadruple mutant (M133L/V203A/N239Y/N268D) was designed, increasing thermal stability by about 5.6 °C.[9] This thermally stabilized mutant, commonly called "T-p53", has become a standard for biophysical experiments.

The mutation Y220C, located in the core domain, is one of the well-known hotspots.[10] The substitution of tyrosine to cysteine thermally destabilizes the core domain by about 8 °C.[11] However, this mutation yields a small, hydrophobic cleft. Early work in targeting this Y220C cavity has yielded a carbazole derivate, PK083 that stabilizes T-p53C-Y220C by

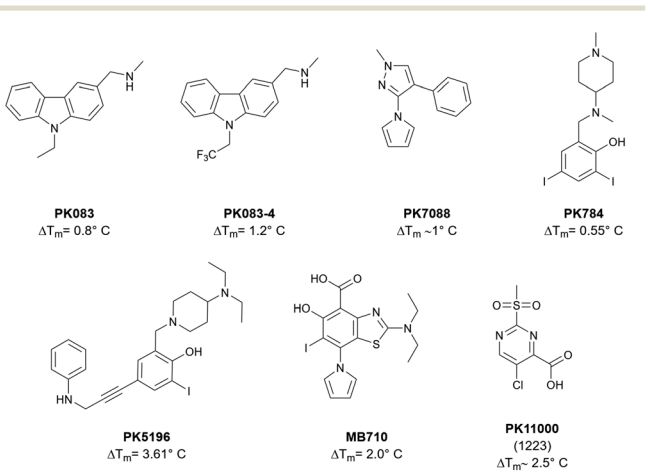

Fig. 1 An exemplary selection of published T-p53C-Y220C binders.[12–17] Compounds PK083 and PK784 occupy the central cavity and subsite 1, while PK5196, PK7088 and MB710 additionally bind subsite 2. The pyrrole group of PK784 and MB710 also enables them to engage to the deep subsite 3. PK11000 showed covalent addition to multiple cysteines.

[a] *Lab for Molecular Design & Pharm. Biophysics, Institute of Pharmaceutical Sciences, Department of Pharmacy and Biochemistry, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany. E-mail: frank.boeckler@uni-tuebingen.de*
[b] *Institute of Pharmaceutical Sciences, Pharmaceutical (Bio-)Analysis, University of Tübingen, Auf der Morgenstelle 8, 72076, Tübingen, Germany*
[c] *Interfaculty Institute of Biochemistry, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 34, 72076 Tübingen, Germany*
[d] *Institute of Organic Chemistry, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 18, 72076 Tübingen, Germany*
[e] *Department of Protein Evolution, Max Planck Institute for Biology Tübingen, Max-Planck-Ring 5, 72076 Tübingen, Germany*
[f] *Interfaculty Institute for Biomedical Informatics (IBMI), Eberhard Karls Universität Tübingen, Sand 14, 72076 Tübingen, Germany*

0.8 °C (125 μM cmpd. conc.).[12] Trifluorination at the *N*-ethyl anchor of PK083 improved the stabilization to 1.2 °C (125 μM cmpd. conc.) by benefiting from multipolar interactions and fluorine–sulfur contacts.[13] This compound and other selected are depicted in Fig. 1.

Quantum chemical analysis of the cavity showed that Leu145 can be addressed by halobenzene moieties, which initially led to the design of a halogen-enriched fragment library.[14] The 2,4-diiodophenol derivative PK784 was identified and crystallized, showing that a halogen bond with Leu145 as postulated is formed. Also, the compound increases the melting temperature by 0.55 °C (250 μM cmpd. conc.). Optimization of this scaffold led to PK5196, stabilizing T-p53C-Y220C by 3.61 °C (250 μM cmpd. conc.) using an acetylene linker to extend into subsite 2, illustrated in Fig. 2.[14]

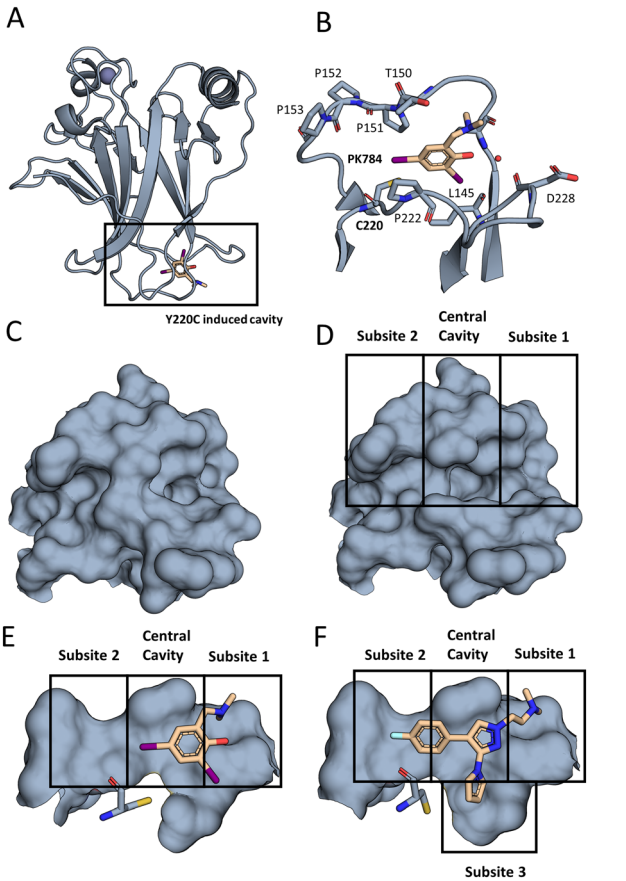

Fig. 2 The Y220C mutation leads to a druggable cleft. (A) Overall view of T-p53C-Y220C bound to PK784 (4AGL[14]). The location of the mutation is highlighted, leading to a druggable cleft. (B) PK784 engages in a halogen bond with L145. The other iodine points towards subsite 2 and can be used as vector for ligand growing. (C) Surface representation of Y220 and surrounding amino acids in T-p53C (1UOL[18]). (D) The mutation opens the central cavity, connecting subsite 1 and subsite 2 (4AGL[14]). (E) The lower part of the cleft is not targeted as C220 points towards the compound. (F) Upon binding of an electron rich group, *e.g.* pyrrole in PK7242, the cysteine is displaced enlarging subsite 3 and enabling targeting (3ZME[15]).

Later efforts identified PK7088 from a fragment library as a Y220C cleft binder, bearing a different scaffold. Interestingly, the pyrrole sidechain of this compound points into a deep internal cavity (subsite 3, Fig. 2) formed by a flip of C220, leading to a stabilization of 1 °C (350 μM cmpd. conc.).[15]

Based on this discovery, the aminobenzothiazole derivative MB710 was synthesized, linking the diiodophenol scaffold to the pyrrolo sidechain. This structure yields a thermal stabilization by 2 °C (250 μM cmpd. conc.).[16]

These promising results were the basis for the design of a diversity-optimized halogen-enriched fragment library (HEFLib).[19] In this generalized HEFLibs approach, aiming for a collection of chemical probes for investigating halogen bonding by fragment-based drug discovery,[20] library design and selection principles became independent from the initial target-focused approach.[14] From kinase drug discovery, we learned that molecular design of halogen bonds at an advanced state of the lead optimization process can be quite challenging and does not necessarily reach its full potential to improve ligand affinities.[21] Thus, we concluded that focusing on the earliest stages of the drug discovery process by generalizing our HEFLib's strategy can also provide the chance to establish unconventional binding modes based on specific halogen bonding motifs and allows to harness the chemotypes of such hits as novel lead structures with great impact on patentability.[19,20] As the design was led by diversity and limited by availability, only 14 of the 191 compounds contain an iodine. This also means that compounds with strongly tuning groups were not specifically selected. The compounds based on PK784 on the other hand contain an iodine and a positive charge, leading to strong tuning effects.[22]

In principle, there is a multitude of suitable electron-rich interaction partners for accepting a halogen bond in a typical binding site:[23] the backbone carbonyl,[24,25] the peptide bond,[26] the sulfur of methionine[27] or cysteine, the nitrogen of histidine,[28] the oxygen of the hydroxyl group of serine, threonine or tyrosine, the carboxamides of asparagine and glutamine, as well as the carboxylates of aspartate and glutamate,[29] and the π-systems of phenylalanine, tyrosine, tryptophan, or histidine.

Based on this abundance of interaction partners and the different construction of the generalized HEFLib, we deemed it an interesting experiment to revisit the challenging binding site of T-p53C-Y220C with such an unbiased library. It is noteworthy that this diversity-optimized HEFLib has been thoroughly tested and characterized, resulting in many different hits on various targets.[30–32]

Another approach for stabilization was identified by covalent modification of cysteines other than C220.[17] A 2-sulfonylpyrimidine was identified that covalently modifies C182 and C277, increasing the T_m up to about 2.5 °C without losing affinity towards DNA. This effect is independent of the Y220C mutation, which makes it a great candidate for a general T-p53C stabilization approach.

As the thermal destabilization of the Y220C mutation is of greatest concern, we performed differential scanning fluorimetry (DSF) as one of our primary screening techniques. Additionally, we used saturation transfer difference (STD) NMR for hit identification. Comparing target binding to its functional results by these two techniques can yield valuable insights for the drug discovery process. All hits found by these independent techniques were validated by ^{1}H,^{15}N-HSQC. Thus, binding modes independent of the Y220C induced cleft were detectable during primary screening as well and were classified through the validation process.

2 Results and discussion

In total, 14 hits were identified by the primary screens, of which ten were confirmed by ^{1}H,^{15}N-HSQC. A compound was considered as a hit if either the STD signal was sufficiently larger than the local background noise or if the ΔT_m was greater than 0.5 °C,[33] measured by DSF. Because it was shown that fragments can also stabilize T-p53C-Y220C through alkylation of surface-exposed cysteines rather than binding in the Y220C cleft,[17] we performed GSH reactivity measurements with each ^{1}H,^{15}N-HSQC hit as a surrogate assay for covalent binding. These measurements were also used for general stability assessment, as it is possible that strongly tuned halogenated fragments can be degraded *via* a S_NAr-type reaction.[30] All hits identified in this study are summarized in Table 1. For clarity and simplicity the HEFLib IDs were reduced to their last four digits.

The DSF results of the triplicate measurement of the STD and DSF hits during the screen are shown in Fig. 3. The compounds that were considered hits are all significant ($p < 0.05$) compared to the reference.

Table 1 All hits identified in this study are listed. A compound was considered as a hit if either the STD signal was present or the ΔT_m was larger than 0.5 °C. The ^{1}H,^{15}N-HSQC indicates if significant peak shifts were observed. The GSH column indicates the measured stability with GSH. Because compound 1151 does not contain a C–H bond, no STD spectrum can be recorded. Only ^{1}H,^{15}N-HSQC hits were measured for GSH stability. For the STD experiments a protein to ligand ratio of 20 μM : 1 mM (1 : 50), for DSF 6 μM : 1 mM (1 : 125) and HSQC 65 μM : 2 mM (1 : 30) was used

Compound	ΔT_m [°C]	STD	HSQC	$t_{1/2}$ GSH [h]
0116	−0.05	Yes	Yes	>100
0403	0.05	Yes	No	—
0404	−0.10	Yes	Yes	>100
0459	−0.05	Yes	Yes	>100
0522	0.30	Yes	No	—
0660	0.10	Yes	No	—
1151	0.80	NA	Yes	>100
1218	−0.20	Yes	Yes	>100
1223	2.0–2.55	Yes	Yes	0.81
1243	−0.25	Yes	Yes	>100
1246	0.15	Yes	No	—
4482	0.5–1.75	No	Yes	2.8
7394	−0.15	Yes	Yes	61
7405	0.20	Yes	Yes	>100

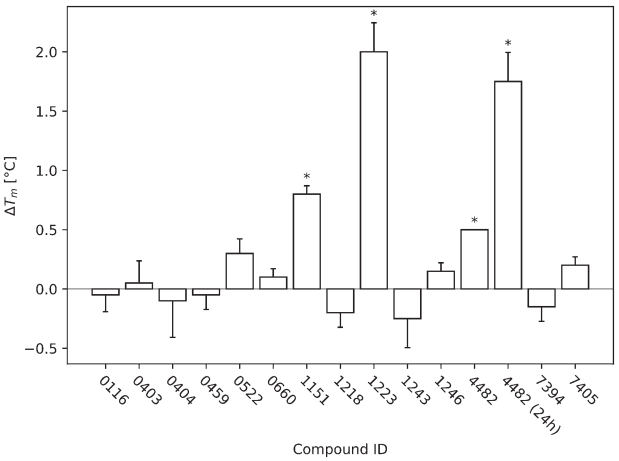

Fig. 3 The DSF screening results of all identified initial hits are displayed. The T-p53C-Y220C concentration was 8 μM and the ligand concentration 1 mM (ratio 1 : 125). The DSF runs were started after 30 min of incubation. All temperature changes that were significant ($p < 0.05$) compared to the reference (T-p53C-Y220C without ligand) of the respective run are marked with an asterisk. During the run of 4482, all three measurements showed exactly the same difference to the reference and therefore no error bar is displayed. Additionally, for 4482 another run with 24 h incubation is displayed, showing time dependent increase in stabilization.

Based on the ^{1}H,^{15}N-HSQC data, hits could be categorized into three groups. By comparing the peak shifts with the ^{1}H,^{15}N-HSQC spectra of crystallographically confirmed Y220C cleft binders,[12,14] one group was identified as the Y220C cleft binding group. Another group was confirmed to bind covalently to T-p53C-Y220C. The third set contained significant peak shifts but does not suggest Y220C cleft binding. The chemical structure and corresponding group are depicted in Fig. 4.

2.1 Targeting the Y220C cleft

Four compounds (0459, 1151, 7394 and 7405) showed at least one peak shift, common to published cleft binders. However, none of the ^{1}H,^{15}N-HSQC hits yielded fittable ITC spectra (data not shown), they contain partially new scaffolds. The compound 7405 contains a *meta*-di-halogen moiety vicinal to an H-bond donor (amine). This compound has similarity to the first structure with confirmed halogen bonding to T-p53C-Y220C (PK784, Fig. 1 and 2). This scaffold can give rise to new synthetic routes and modifications.

Of the proposed cleft-binding compounds, only compound 1151 stabilized the protein by more than 0.5 °C. The ^{1}H,^{15}N-HSQC spectrum (Fig. 5A) shows multiple peak shifts indicating cleft binding. The DSF curves with addition of 1151 (ratio 1 : 125), shown in Fig. 5B, are shifted by 0.8 °C.

Compounds 0459 and 1151 both contain a pyrazole scaffold with a halogen in position 4. A structure was solved of T-p53C-Y220C bound to 1151 (PDB: 8A92). Data collection and refinement statistics are displayed in Table 2. The structure, especially in chain A, indicates multiple binding

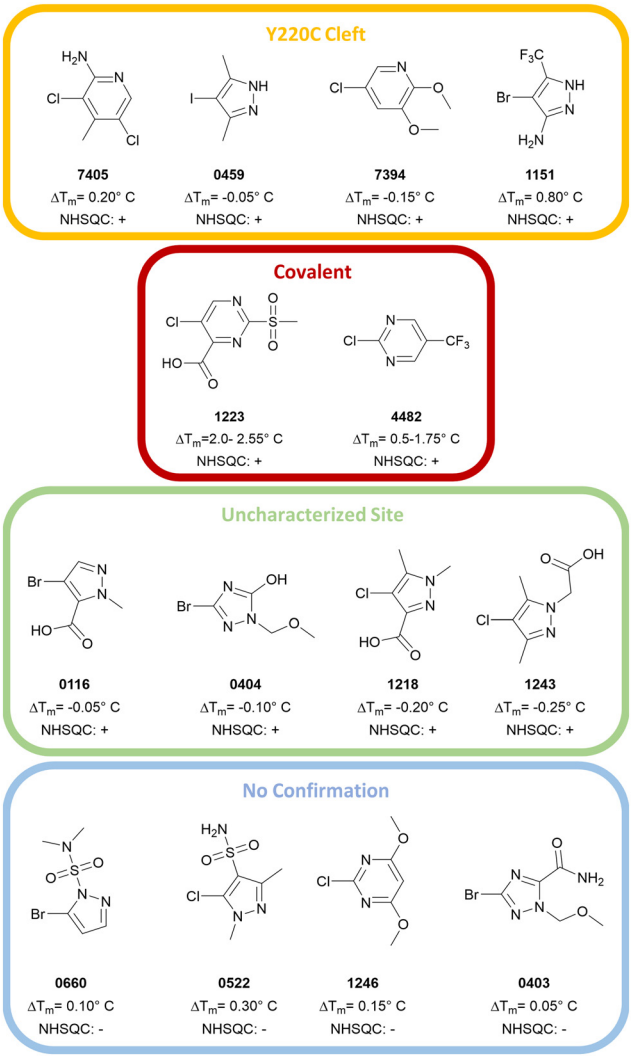

Fig. 4 Overview of the hits from the primary screens. Hits confirmed by ^{1}H,^{15}N-HSQC were categorized as either Y220C cleft binders (0459, 1151, 7394, 7405), covalent binders (1223/PK11000, 4482) or binders to an uncharacterized site (0116, 0404, 1218, 1243). The compounds 0660, 0522, 1246 and 0403 could not be confirmed by ^{1}H,^{15}N-HSQC.

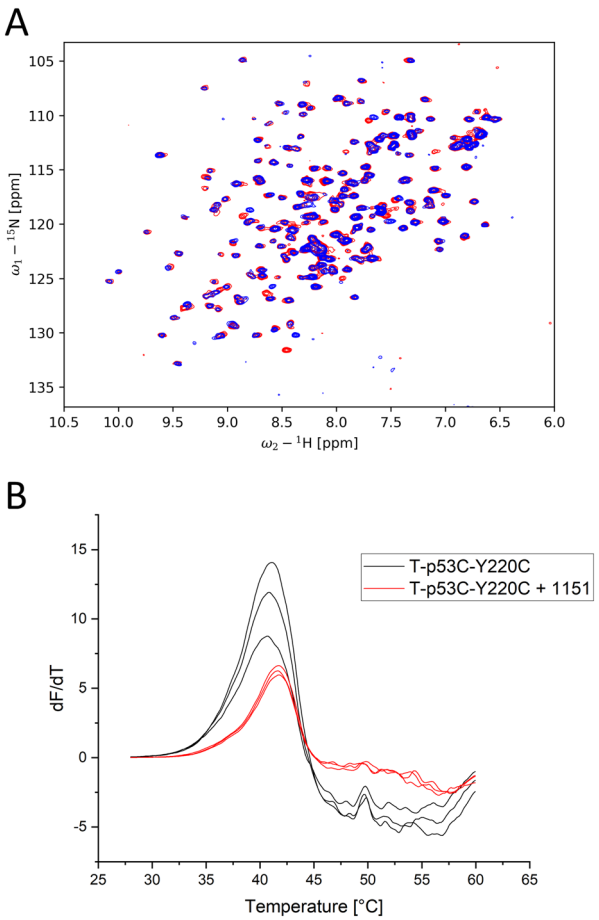

Fig. 5 (A) ^{1}H,^{15}N-HSQC of T-p53C-Y220C with 1151 at 2 mM (blue) and without addition of a ligand (red). Multiple peak shifts can be observed, all of which indicate Y220C cleft binding. (B) First derivative of the melting curves of T-p53C-Y220C and T-p53C-Y220C with 1151 of the initial screen (30 min incubation).

poses, which could be identified by the anomalous signal of the bromine (Fig. 6). For chain B the occupancy was low and only one pose was built. The poses show that the CF_3 group points towards the cleft and the bromine pointing towards L145 or the subsite 2. The electron-rich CF_3 group displaces C220, opening subsite 3. This binding pose is similar to the ones containing a pyrrole ring, *e.g.* MB710 (Fig. 1 and 2). Even though the compound has a low K_D value (>1 mM, based on non-fittable ITC data) it stabilizes T-p53C-Y220C by about 0.8 °C. It is likely that the binding is driven by hydrophobic interactions of the CF_3 group to the cleft. No electron density could be observed for the amine.

In one pose the bromine points towards the backbone oxygen of L145 with a distance of 3.0 Å and an angle of 153.4°. Based on this pose we calculated the adduct-formation energy of the compound and *N*-methylacetamide as a backbone model. In order to estimate the effect of the CF_3-group we calculated the V_{max} at an isodensity level of 0.02 au and the adduct-formation energy (ΔE) of 1151 and closely related scaffolds at the MP2/TZVPP level of theory. The results are displayed in Fig. 7. In addition, this bromine accepts an orthogonal hydrogen bond from the backbone nitrogen of T230.

The CF_3 group is the main contributor for the tuning of the system, increasing the V_{max} from 0.1597 au to 0.1737 au, comparable with iodobenzene (0.182 au).[34] As the amine only has a small effect on tuning, the vector could be used for fragment growing. Another approach could be converting the amine to a tertiary amine, adding a positive charge and once more strongly tune the system.[22]

2.2 Covalent modification

Because the 2-sulfonylpyrimidine PK11000 (Fig. 1) was added to the HEFLib with the ID 1223, a positive control for covalent modification was present. The glutathione assays confirmed that each fragment measured was stable with a

Table 2 Data collection and refinement statistics of the dataset of T-p53C-Y220C soaked with compound 1151 (PDB: 8A92). Values in parentheses indicate the respective value of the highest resolution bin

	Compound 1151 (PDB: 8A92)
Wavelength [Å]	0.92
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* [Å]	65.1, 74.1, 105.2
α, *β*, *γ* [°]	90, 90, 90
Resolution range [Å]	50–1.37 (1.45–1.37)
Redundancy	13.2 (13.3)
Completeness [%]	100 (100)
Mean $I/\sigma(I)$	14.4 (1.0)
R-meas [%]	11.7 (256.8)
CC1/2 [%]	99.9 (49.7)
Wilson *B* [Å^2]	22.8
Resolution included [Å]	43.68–1.37
R_{work}/R_{free} [%]	15.32/18.32
Bond RMSD [Å]	0.008
Angle RMSD [°]	0.99
Ramachandran [%] (favored/allowed/outliers)	98.98/1.02/0
Rotamer outliers [%]	1.43
All-atom Clashscore	1.44
Average *B* factor [Å^2]	
Overall	22.34
Protein	20.03
Ligand	41.4
Water	35.58

half-life greater than 20 h, with the exception of 4482 and the positive control as aforementioned. A half-life of 2.8 h was observed for compound 4482, which indicated a covalent reaction with the protein. The ^{1}H,^{15}N-HSQC, as displayed in Fig. 8, shows large peak shifts, similar to PK11000, also suggesting covalent modification. We were able to confirm covalent binding by mass spectrometry. The deconvoluted ESI-MS spectrum (Fig. 8, panel C) indicates that after 4 h of incubation with a molar protein-to-compound ratio of 1 : 125 up to three molecules of 4482 were bound to T-p53C-Y220C. Among them, the species with two attached molecules showed highest intensity. Unmodified protein was not detectable. In contrast, the MS spectrum of the triple mutant C124/182/277S, shows only the species with one bound fragment. Unmodified protein could be identified. This indicates that covalent binding of 4482 to the now absent cysteines occurs. Using DSF measurements, a time-dependent stabilization could be identified (Fig. 8). The ΔT_m of T-p53C-Y220C increased from 0.9 °C after 4 h incubation up to 1.75 °C after a 24 h incubation period.

In comparison, PK11000 resulted in a larger stabilization of T-p53C-Y220C up to a maximum ΔT_m of about 2.5 °C. In contrast to 4482, a maximum number of two modified cysteines was evident in the ESI-MS spectra of PK11000. This highlights the selectivity of PK11000 for specific p53 cysteines.[17]

Interestingly, Bauer *et al.*[17] could not detect any stabilization by the compound 2-chloro-5-(fluoromethyl)-pyrimidine. We measured significant stabilization with 4482, the trifluoro analog of the compound. This emphasizes the need of the CF_3 moiety as a strong electron-withdrawing group (EWG) for covalent binding. In general, compound 4482 contains a similar scaffold to the known covalent modifier PK11000. Both are pyrimidines with leaving groups in position 2 and additional EWGs to enhance S_NAr reactivity. These structural similarities should explain their comparable behavior towards T-p53C and T-p53C-Y220C. Moreover, it demonstrates the potential of their common scaffold as a covalent, Y220C-cleft independent, rescue for the oncogenic p53-Y220C mutant.

2.3 Identification of an uncharacterized site

Compounds 0116, 0404, 1218, 1243 and did not show typical peak shifts of previously described peaks. Examples of these peak shifts are displayed in Fig. 9. As there is no structural confirmation of another binding site, it is difficult to interpret the reason for the peak shift. The lack

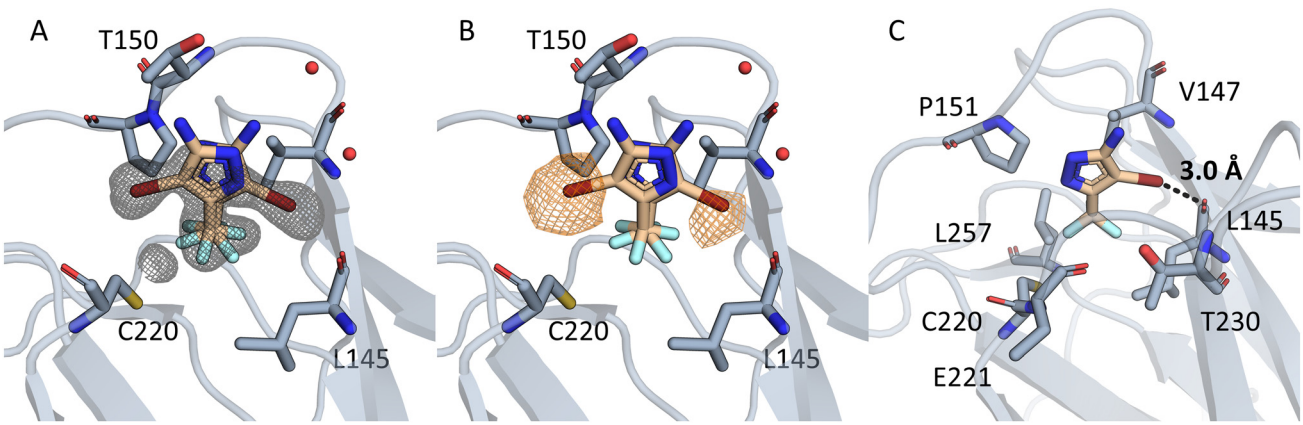

Fig. 6 (A) Electron density is displayed in grey around the compound 1151 as an unbiased omit map contoured at 3σ. In chain A two binding poses can be seen. In both poses the CF_3 group points towards C220 and engages in hydrophobic contacts. (B) The unbiased anomalous difference map is contoured at 4.5σ in orange. The signal shows that the bromine is located in two positions, but not towards C220. (C) The halogen engages in a halogen bond, with a distance of 3.0 Å and an angle of 153.4°. Residues in 4.5Å around the CF_3 group are displayed.

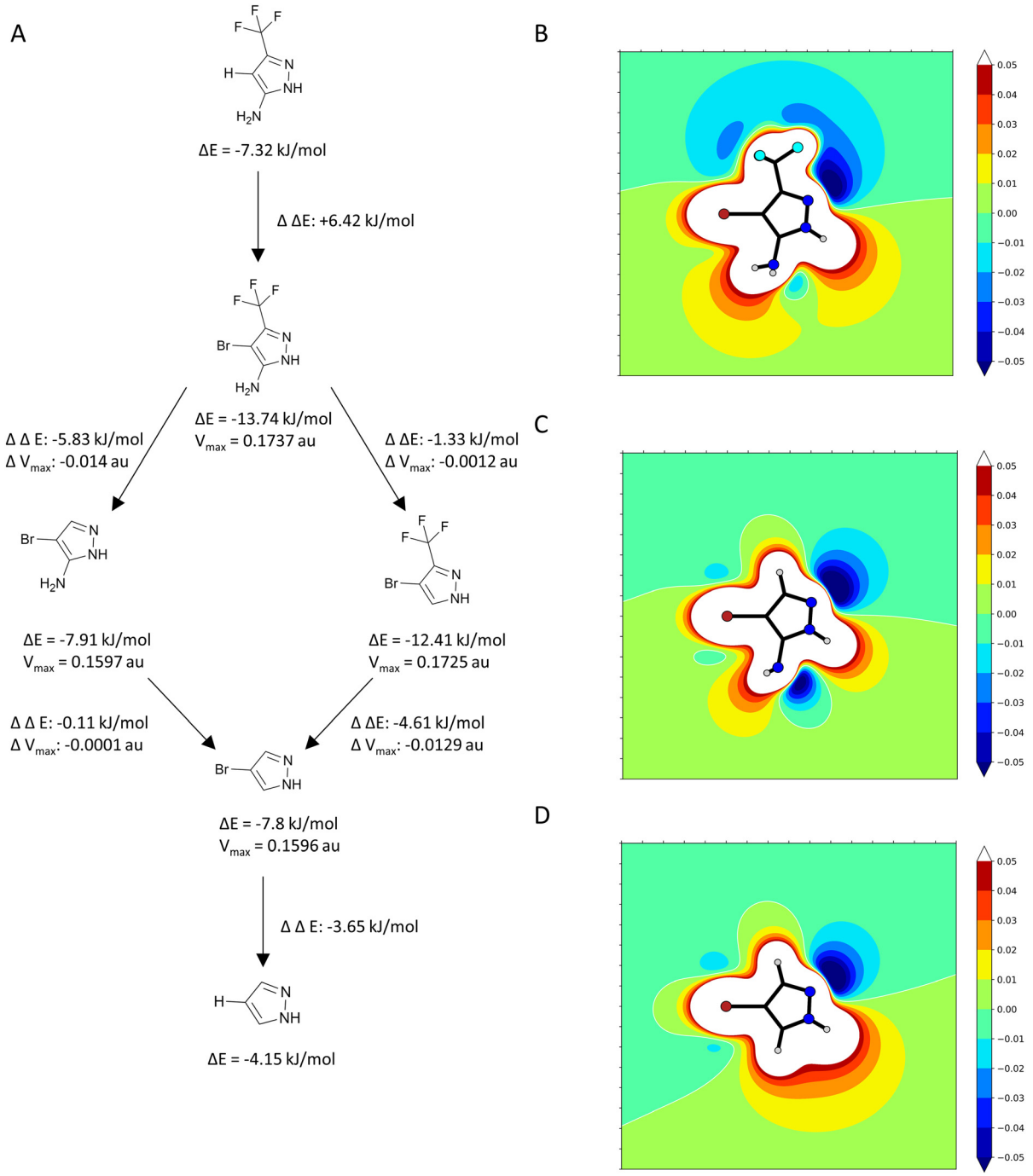

Fig. 7 (A) Flowchart of the effects on adduct-formation energy (ΔE) and V_{max} by systematically adding or removing substituents based on 1151. The CF_3 group has the largest effect on ΔE and V_{max}. (B) ESP plot of 1151. (C) ESP plot of 1151 after removing the CF_3 group. (D) ESP plot of 1151 after removing the CF_3 and NH_2 group.

of a deep cavity or cleft makes targeting of p53C very difficult. All compounds with the exception of 0404 contain a carboxyl group, which could engage in loose binding to the positively charged DNA binding interface of p53C and lead to the aforementioned peak shifts. Comparing the two prominent shifts (ω_1 = 130.29 ppm, ω_2 = 8.8 ppm; ω_1 = 110.9 ppm, ω_2 = 8.73 ppm) to the spectra of PK11000 or 4482, similar shifts can be observed. This could indicate, that the mostly uncharacterized binding site is near any of the alkylated cysteines.

3 Conclusion

The lack of specific cancer treatment for p53-Y220C and high amount of annual cases of an estimated 130 000 underlines the importance of continuous drug discovery for this

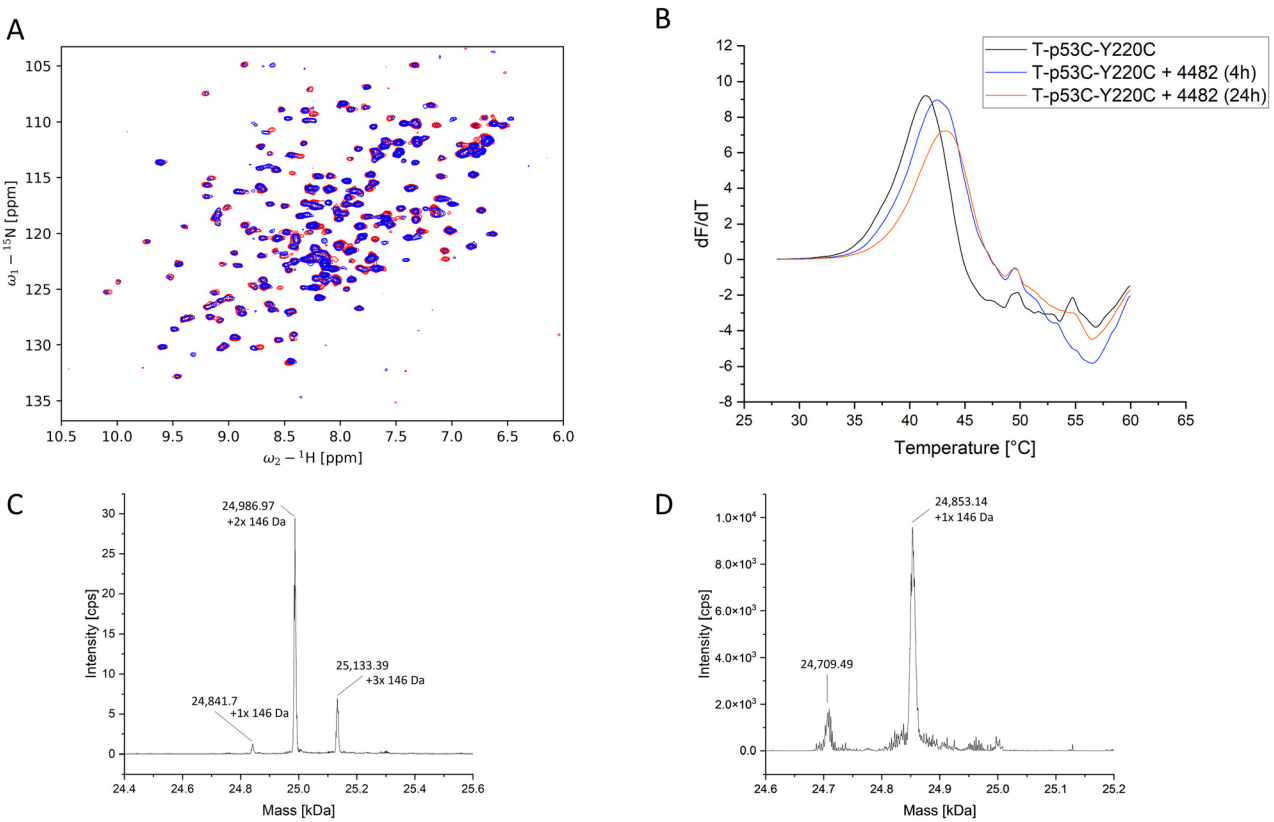

Fig. 8 (A) ^{1}H,^{15}N-HSQC of T-p53C-Y220C without compound (red) and 2 mM 4482 (blue). Multiple prominent peak shifts can be observed. (B) First derivative of the melting curve of T-p53C-Y220C and T-p53C-Y220C with 4482 after 4 h and 24 h incubation. (C) Deconvoluted ESI-MS spectrum of T-p53C-Y220C (50 μM) with 6.25 mM 4882 after 4 h incubation at 20 °C. The shift of the peaks is about 146 Da, which corresponds to the size of the attached 4482. (D) Deconvoluted ESI-MS spectrum of T-p53C-C124/182/277S (50 μM) with 6.25 mM 4482 after 4 h incubation at 20 °C. The major peak corresponds to the single alkylated protein.

target.[35–38] The first confirmation of halogen bonding in T-p53C-Y220C led to the development of the diversity-optimized HEFLib. A total of ten hits from this diversity-optimized library could be confirmed by ^{1}H,^{15}N-HSQC. The development of a V_{max}-optimized HEFLib could lead to an improvement of the hit rate.[30] The diversity-optimized HEFLib did not yield another low Millimolar binding compound, but identified a multitude of diverse scaffolds binding to the protein. These scaffolds can now be further exploited by optimizing target binding and more importantly stabilization.

The HEFLib additionally produced more unconventional binding modes, by containing a covalent binder. This modification is independent of the Y220C mutation and could potentially lead to general purpose stabilizers with much broader applicability in the rescue of destabilized p53 cancer mutants. Another aim could be modifying these scaffolds to reestablish or modulate DNA binding in hot-spot mutations which loose their function based on altered DNA contacts.

Many of the identified and confirmed hits did not interact in the Y220C mutation-induced cavity as we had anticipated. For these hits further studies are needed to reveal their binding site and its possible implication for altering or rescuing p53 function. It should be noted that besides loss-of-function mutations, there are also oncogenic gain-of-function mutations known for p53.[39–41] Thus, new binding modes could provide avenues for therapeutic intervention for both aspects of altered p53 function. Of course, tractability and druggability need to be shown for such a new binding site.[42]

The compound 1151 could be further characterized and used as a spy molecule for FAXS NMR.[43,44] This could further facilitate screening experiments for the identification of novel fragments binding to the Y220C-induced cleft.

4 Materials and methods

4.1 Molecular biology

In general, expression and purification was performed as previously described.[16] The plasmid was cloned into BL21pLysS cells and the protein was expressed over night at 24 °C. The lysed sample in lysis buffer (KP_i 50 mM, NaCl 300 mM, imidazole 10 mM, TCEP 2 mM, pH = 8) was loaded onto a NiNTA column (Cytiva) and eluted using a gradient from 0–100% elution buffer (KP_i 50 mM, NaCl 300 mM, imidazole 250 mM, TCEP 2 mM, pH = 8). Then, the tag was cleaved over night using the tobacco etch virus protease. The sample was diluted 8-fold with heparin buffer A (KP_i 25 mM, NaCl 0 mM, DTT 5 mM, pH = 7.5) before it was loaded onto a HiTrap

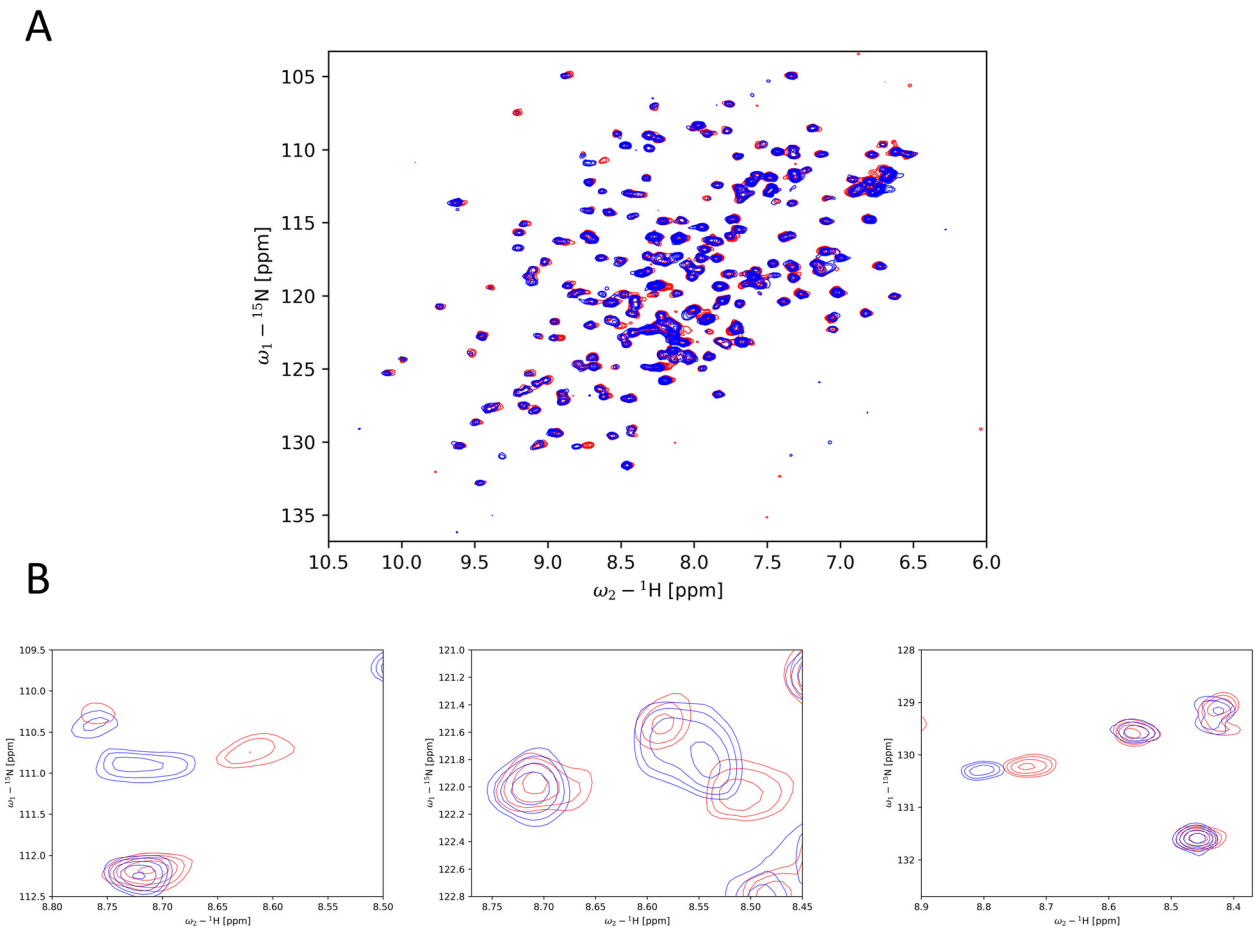

Fig. 9 Example of peak shifts of an uncharacterized binding site. (A) Overlay of peak shifts shown by compound 0116 at 2 mM (blue) in comparison to reference (no ligand) (red). (B) Typical peak shifts observed in the subgroup of the binders to an uncharacterized binding site.

Heparin HP (GE Healthcare) column. The sample was eluted with 40% heparin buffer B (KP_i 25 mM, NaCl 2 M, DTT 5 mM, pH = 7.5). A final size exclusion chromatography step using a HiLoad 26/60 Superdex 75 pg (GE Healthcare) was performed (KP_i 25 mM, NaCl 150 mM, DTT 5 mM, pH = 7.2). Fractions containing protein were finally pooled and flash frozen in liquid nitrogen.

For labelled expression, M9 minimal medium[45] was used supplemented with 1 g L^{-1} of $^{15}NH_4Cl$ prior to inoculation.

For mutagenesis, a pET24a-HLT vector with the thermostable p53 core domain (94–312) served as template. The Q5 site-directed mutagenesis kit (New England BioLabs) was used to produce the triple cysteine mutant C124/182/277S.

All constructs used, are displayed in Table 3.

4.2 Differential scanning fluorimetry (DSF)

The melting temperature of T-p53C-Y220C in the presence or absence of fragments was determined by differential scanning fluorimetry (DSF). DSF measurements were performed on a Qiagen Rotor-Q Model-5-Plex HRM real-time PCR instrument using SYPRO Orange as fluorescent dye (final concentration 5×). 8 μM protein in phosphate buffer (25 mM KPi, 150 mM NaCl, 1 mM TCEP, pH 7.2, 5% DMSO [v/v]), and a final compound concentration of 1 mM were used.[17] The temperature was ramped from 28 °C to 60 °C with a heating rate of 270 °C h^{-1}.[14] Excitation and emission filters were set to 490 and 580 nm. For time-dependent DSF measurements, the samples were incubated at 20 °C on a rotating shaker. The melting temperature of T-p53C-Y220C with or without compounds was determined from the maximum of the first derivative of the melting curve using OriginPro2020. All measurements were performed in triplicates. ΔT_m was calculated by subtracting the resulting T_m of T-p53C-Y220C from the T_m of the compound samples. A temperature increase of at least 0.5 °C was defined as a parameter for a fragment hit.[33]

4.3 Saturation transfer difference (STD) NMR

The STD experiments were performed as previously described, using the same mixtures.[30] The spectra were recorded on a Bruker Avance III HDX 700 with a 5 mm Prodigy TCI cryo probehead. Compounds were considered as a hit, if the signal was sufficiently larger than the local background noise.

Table 3 All sequences used are displayed. The biophysical screens are DSF, STD, ESI-MS and ^{1}H,^{15}N-HSQC. For crystallization a construct lacking the linker GGS was used

Construct	Sequence (N′–C′)
T-p53C-Y220C (biophysical screens)	GGSSS SVPSQ KTYQG SYGFR LGFLH SGTAK SVTCT YSPAL NKLFC QLAKT CPVQL WVDST PPPGT RVRAM AIYKQ SQHMT EVVRR CPHHE RCSDS DGLAP PQHLI RVEGN LRAEY LDDRN TFRHS VVVPC EPPEV GSDCT TIHYN YMCYS SCMGG MNRRP ILTII TLEDS SGNLL GRDSF EVRVC ACPGR DRRTE EENLR KKGEP HHELP PGSTK RALPN NT
T-p53C-Y220C (crystallization)	SSSVP SQKTY QGSYG FRLGF LHSGT AKSVT CTYSP ALNKL FCQLA KTCPV QLWVD STPPP GTRVR AMAIY KQSQH MTEVV RRCPH HERCS DSDGL APPQH LIRVE GNLRA EYLDD RNTFR HSVVV PCEPP EVGSD CTTIH YNYMC YSSCM GGMNR RPILT IITLE DSSGN LLGRD SFEVR VCACP GRDRR TEEEN LRKKG EPHHE LPPGS TKRAL PNNT
T-p53C-Y220C (C124/182/277S)	GGSSS SVPSQ KTYQG SYGFR LGFLH SGTAK SVTST YSPAL NKLFC QLAKT CPVQL WVDST PPPGT RVRAM AIYKQ SQHMT EVVRR CPHHE RSSDS DGLAP PQHLI RVEGN LRAEY LDDRN TFRHS VVVPC EPPEV GSDCT TIHYN YMCYS SCMGG MNRRP ILTII TLEDS SGNLL GRDSF EVRVC ASPGR DRRTE EENLR KKGEP HHELP PGSTK RALPN NT

4.4 ^{1}H,^{15}N-Heteronuclear single quantum coherence NMR

Spectra were recorded on an 800 MHz Bruker Avance-III at 293 K using a final protein concentration of 65 μM (5% DMSO-d6 [v/v]) and a final ligand concentration of 2 mM.

In total, 1024 data points were collected in the direct dimension (^{1}H) and 128 in the indirect dimension (^{15}N). Data processing was performed using Bruker Topspin 4.1 and analysis using NMRFAM-SPARKY 3.19.[46] Peak shifts were considered significant, if

$$\Delta\delta\left(^{1}\mathrm{H}/^{15}\mathrm{N}\right) = \sqrt{\left(\Delta\delta(^{1}\mathrm{H})\right)^{2} + \left(\frac{\Delta\delta(^{15}\mathrm{N})}{5}\right)^{2}} \quad (1)$$

was larger than 0.04 ppm.[14,47] All NMR figures were prepared using the nmrglue Python package.[48]

4.5 Glutathione assay

GSH stability studies were performed according to a protocol established for heterocyclic electrophilic fragments by Keeley *et al.*[49]

Reaction conditions were PBS buffer pH 7.4, 10% acetonitrile, 100 μM ketoprofen as an internal standard, 250 μM fragment, and 5 mM GSH excess at 37 °C. Measurement times were after 0, 1, 2, 4, 8, 12, and 24 h. For very reactive fragments ($t_{1/2} < 5$ h) analysis was performed every 20 min. The mixture was analyzed by HPLC with UV-detection. The reaction of the compounds was detected by measuring the decreasing area under the curve (AUC) of the fragment relative to the internal standard. The declining AUC was fitted to pseudo-first order kinetics and $t_{1/2}$ was calculated using the following equation:

$$t_{1/2} = \ln\left(\frac{2}{k}\right) \quad (2)$$

Measurements were performed as duplicates with GSH. In addition, single measurements without GSH in PBS buffer were carried out for each fragment to check for hydrolytic degradation.

4.6 Mass spectrometry

T-p53C-Y220C or T-p53C-C124/182/277S (50 μM) in phosphate buffer (25 mM KP_i, 150 mM NaCl, 1 mM DTT, pH 7.2) were incubated with 125 mM compound dissolved in DMSO, yielding a final concentration of 5% [v/v] DMSO and 6.25 mM protein with a protein to compound ratio of 1 : 125. The mixture was incubated for 4 h at 20 °C while shaking.[17]

The UHPLC-system consisted of an Agilent (Waldbronn, Germany) 1290 Infinity binary pump (G4220A) and a thermostated column compartment (G1316C). Between the column and the ion source a Valco EHMA diverter valve (2-pos/6-port) from VICI (Schenkon, Switzerland) was installed for online de-salting of the samples. Mobile phase A was water + 0.1% (v/v) formic acid and mobile phase B was ACN + 0.1% (v/v) formic acid. The flow rate was set to 0.5 mL min^{-1} with the following gradient: 0–2 min: 5% B (de-salting of sample, flow to waste), 2–10 min: 5–80% B, 10–12 min: 80% B, 12–17 min: 5% B. The column temperature was set to 50 °C. A TripleTOF 5600+ mass spectrometer from Sciex (Darmstadt, Germany) was used with a Duospray ion source (ESI interface) in positive ionization mode. The following MS instrument parameters were used: curtain gas (CUR): 35 psi, nebulizing gas (GS1): 50 psi, heater gas (GS2): 40 psi, ion spray voltage floating: 5100 V, source temperature: 550 °C, collision energy (CE): 30 V, declustering potential (DP): 220 V. The mass range in TOF MS mode was set from 500 to 5000 m/z with an accumulation time of 500 ms. The IntactProteinMode script from Sciex was used to optimize advanced MS settings for protein analysis. Data acquisition was performed with Analyst TF 1.8.1 software (Sciex). Data analysis was performed using PeakView software 2.2.0 (Sciex) using the BioToolKit (2.2.0.) for deconvolution of the mass spectra.

4.7 Crystallization and data collection

The protein was concentrated to 5 mg mL^{-1} and mixed 1 : 1 with reservoir solution (100 mM HEPES (pH = 7.15), 19% PEG4000 and 10 mM DTT) using the sitting drop vapor diffusion technique and by performing streak seeding. Crystals grew within a few days. For soaking, crystals were transferred in a 50 mM or saturated compound solution in a cryo-protectant buffer (reservoir solution and an additional 20% glycerol) over night.

Data sets were obtained at the Swiss Light Source (SLS) (Villigen, Switzerland) at the X06DA (PXIII) beamline, using a Pilatus-2 M-F detector.

Data processing and reduction was performed using XDS.[50] To obtain initial phases by molecular replacement, 4AGL was used as a search model for PHASER included in the CCP4 suite.[14,51,52] Structure and phase improvement was performed using multiple cycles of manual model building in Coot and structure refinement using PHENIX.[53,54] The anomalous difference map was generated using CAD and fft, both from the CCP4 suite.[55–57] Ligand restraints were generated using AceDRG.[58,59]

4.8 V_{max} and adduct formation energy calculations

4.8.1 MP2 structure optimizations and single point calculations. Geometry optimizations and single point calculations were carried out using TURBOMOLE 7.4.1.[60] A triple-ζ basis set (def2-TZVPP) was used throughout the study.[61] MP2 calculations were done in combination with the resolution of identity (RI) technique and the frozen core approximation.[61–64] The frozen core orbitals were attributed by the default setting in TURBOMOLE by which all orbitals possessing energies below 3.0 au are considered as core orbitals. The SCF convergence criterion was increased to 10^{-8} Hartree for all calculations. Heavy atoms of the model systems were kept frozen during optimization.

4.8.2 Adduct formation calculations. The ligand and the halogen bond accepting moiety (represented as *N*-methylacetamide) were optimized using MP2/TZVPP. Heavy atoms were kept frozen during optimization. For the calculations of the putative adduct formations, the modified ligands were freely optimized using MP2/TZVPP-level of theory. The ligands were subsequently matched onto the geometry of the crystal structure using three atoms of the ligand. The complex formation energies were calculated through single points as the difference between the complex and the sum of the two separate molecules.

4.8.3 2D electrostatic potential (ESP) plots and V_{max} calculations. All ligands were oriented after their geometry optimization employing an in-house script by placing the halogen atom into the negative *X*-axis and the respective ring system into the *XZ*-plane. Electrostatic potentials were calculated using TURBOMOLE 7.4.1 on a regular 2D grid. Visualization was done using a custom Python script. The V_{max} values were extracted at 0.02 au electron isodensity.

Author contributions

Conceptualization was performed by FMB. JS and MC carried out and analyzed ^{1}H,^{15}N-HSQC experiments. MD and MK performed the STD experiments, JS and MD analyzed them. TK performed and analyzed the DSF studies. MS performed the GSH assays. SJ and ML performed and analyzed the ESI-MS experiments. Mutagenesis was done by LE. Crystallographic data reduction and refinement was performed by JS, MB and TS. FMB, JS, TK and MS wrote the final manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We thank Georg Zocher for his help and advice in crystallography. We also thank Andreas Joerger for his expertise in crystallizing T-p53C-Y220C. We gratefully acknowledge the Swiss Light Source (SLS, Switzerland) for beam time.

Notes and references

1 D. P. Lane, *Nature*, 1992, **358**, 15–16.
2 B. Vogelstein, D. Lane and A. J. Levine, *Nature*, 2000, **408**, 307–310.
3 A. C. Joerger and A. R. Fersht, *Annu. Rev. Biochem.*, 2008, **77**, 557–582.
4 A. C. Joerger and A. R. Fersht, *Cold Spring Harbor Perspect. Biol.*, 2010, **2**, a000919.
5 C. J. Brown, S. Lain, C. S. Verma, A. R. Fersht and D. P. Lane, *Nat. Rev. Cancer*, 2009, **9**, 862–873.
6 K. G. Wiman, *Oncogene*, 2010, **29**, 4245–4252.
7 G. M. Popowicz, A. Dömling and T. A. Holak, *Angew. Chem., Int. Ed.*, 2011, **50**, 2680–2688.
8 A. Petitjean, E. Mathe, S. Kato, C. Ishioka, S. V. Tavtigian, P. Hainaut and M. Olivier, *Hum. Mutat.*, 2007, **28**, 622–629.
9 P. V. Nikolova, J. Henckel, D. P. Lane and A. R. Fersht, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 14675–14680.
10 L. Bouaoun, D. Sonkin, M. Ardin, M. Hollstein, G. Byrnes, J. Zavadil and M. Olivier, *Hum. Mutat.*, 2016, **37**, 865–876.
11 M. R. Bauer, A. Krämer, G. Settanni, R. N. Jones, X. Ni, R. Khan Tareque, A. R. Fersht, J. Spencer and A. C. Joerger, *ACS Chem. Biol.*, 2020, **15**, 657–668.
12 F. M. Boeckler, A. C. Joerger, G. Jaggi, T. J. Rutherford, D. B. Veprintsev and A. R. Fersht, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 10360–10365.
13 M. R. Bauer, R. N. Jones, M. G. J. Baud, R. Wilcken, F. M. Boeckler, A. R. Fersht, A. C. Joerger and J. Spencer, *ACS Chem. Biol.*, 2016, **11**, 2265–2274.
14 R. Wilcken, X. Liu, M. O. Zimmermann, T. J. Rutherford, A. R. Fersht, A. C. Joerger and F. M. Boeckler, *J. Am. Chem. Soc.*, 2012, **134**, 6810–6818.
15 X. Liu, R. Wilcken, A. C. Joerger, I. S. Chuckowree, J. Amin, J. Spencer and A. R. Fersht, *Nucleic Acids Res.*, 2013, **41**, 6034–6044.
16 M. G. Baud, M. R. Bauer, L. Verduci, F. A. Dingler, K. J. Patel, D. Horil Roy, A. C. Joerger and A. R. Fersht, *Eur. J. Med. Chem.*, 2018, **152**, 101–114.
17 M. R. Bauer, A. C. Joerger and A. R. Fersht, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, E5271–E5280.
18 A. C. Joerger, M. D. Allen and A. R. Fersht, *J. Biol. Chem.*, 2004, **279**, 1291–1296.
19 J. Heidrich, L. E. Sperl and F. M. Boeckler, *Front. Chem.*, 2019, **7**, 9.

20 M. O. Zimmermann, A. Lange, R. Wilcken, M. B. Cieslik, T. E. Exner, A. C. Joerger, P. Koch and F. M. Boeckler, *Future Med. Chem.*, 2014, **6**, 617–639.

21 A. Lange, M. Günther, F. M. Büttner, M. O. Zimmermann, J. Heidrich, S. Hennig, S. Zahn, C. Schall, A. Sievers-Engler, F. Ansideri, P. Koch, M. Laemmerhofer, T. Stehle, S. A. Laufer and F. M. Boeckler, *J. Am. Chem. Soc.*, 2015, **137**, 14640–14652.

22 R. Wilcken, M. O. Zimmermann, M. R. Bauer, T. J. Rutherford, A. R. Fersht, A. C. Joerger and F. M. Boeckler, *ACS Chem. Biol.*, 2015, **10**, 2725–2732.

23 R. Wilcken, M. O. Zimmermann, A. Lange, A. C. Joerger and F. M. Boeckler, *J. Med. Chem.*, 2013, **56**, 1363–1388.

24 R. Wilcken, M. O. Zimmermann, A. Lange, S. Zahn and F. M. Boeckler, *J. Comput.-Aided Mol. Des.*, 2012, **26**, 935–945.

25 M. O. Zimmermann, A. Lange and F. M. Boeckler, *J. Chem. Inf. Model.*, 2015, **55**, 687–699.

26 M. O. Zimmermann and F. M. Boeckler, *Med. Chem. Commun.*, 2016, **7**, 500–505.

27 R. Wilcken, M. O. Zimmermann, A. Lange, S. Zahn, B. Kirchner and F. M. Boeckler, *J. Chem. Theory Comput.*, 2011, **7**, 2307–2315.

28 A. Lange, M. O. Zimmermann, R. Wilcken, S. Zahn and F. M. Boeckler, *J. Chem. Inf. Model.*, 2013, **53**, 3178–3189.

29 M. O. Zimmermann, A. Lange, S. Zahn, T. E. Exner and F. M. Boeckler, *J. Chem. Inf. Model.*, 2016, **56**, 1373–1383.

30 M. Dammann, M. Kramer, M. O. Zimmermann and F. M. Boeckler, *Front. Chem.*, 2022, **9**, 815567.

31 S. Della Volpe, P. Linciano, R. Listro, E. Tumminelli, M. Amadio, I. Bonomo, W. Elgaher, S. Adam, A. Hirsch, F. Boeckler, F. Vasile, D. Rossi and S. Collina, *Bioorg. Chem.*, 2021, **116**, 105305.

32 M. Dammann, J. Stahlecker, M. O. Zimmermann, T. Klett, K. Rotzinger, M. Kramer, M. Coles, T. Stehle and F. M. Boeckler, *J. Med. Chem.*, in press.

33 J. L. Kaar, N. Basse, A. C. Joerger, E. Stephens, T. J. Rutherford and A. R. Fersht, *Protein Sci.*, 2010, **19**, 2267–2278.

34 A. Lange, J. Heidrich, M. O. Zimmermann, T. E. Exner and F. M. Boeckler, *J. Chem. Inf. Model.*, 2019, **59**, 885–894.

35 K. C. de Andrade, E. E. Lee, E. M. Tookmanian, C. A. Kesserwan, J. J. Manfredi, J. N. Hatton, J. K. Loukissas, J. Zavadil, L. Zhou, M. Olivier, M. N. Frone, O. Shahzada, W. J. R. Longabaugh, C. P. Kratz, D. Malkin, P. Hainaut and S. A. Savage, *Cell Death Differ.*, 2022, **29**, 1071–1073.

36 H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal and F. Bray, *Ca-Cancer J. Clin.*, 2021, **71**, 209–249.

37 J. Ferlay, M. Colombet, I. Soerjomataram, D. M. Parkin, M. Piñeros, A. Znaor and F. Bray, *Int. J. Cancer*, 2021, **149**, 778–789.

38 C. Kandoth, M. D. McLellan, F. Vandin, K. Ye, B. Niu, C. Lu, M. Xie, Q. Zhang, J. F. McMichael, M. A. Wyczalkowski, M. D. M. Leiserson, C. A. Miller, J. S. Welch, M. J. Walter, M. C. Wendl, T. J. Ley, R. K. Wilson, B. J. Raphael and L. Ding, *Nature*, 2013, **502**, 333–339.

39 G. Blandino, A. J. Levine and M. Oren, *Oncogene*, 1999, **18**, 477–485.

40 M. Oren and V. Rotter, *Cold Spring Harbor Perspect. Biol.*, 2010, **2**, a001107.

41 E. Alvarado-Ortiz, K. G. de la Cruz-López, J. Becerril-Rico, M. A. Sarabia-Sánchez, E. Ortiz-Sánchez and A. García-Carrancá, *Front. Cell Dev. Biol.*, 2021, **8**, 607670.

42 K. K. Brown, M. M. Hann, A. S. Lakdawala, R. Santos, P. J. Thomas and K. Todd, *Med. Chem. Commun.*, 2018, **9**, 606–613.

43 C. Dalvit, P. E. Fagerness, D. T. A. Hadden, R. W. Sarver and B. J. Stockman, *J. Am. Chem. Soc.*, 2003, **125**, 7696–7703.

44 C. Dalvit and A. Vulpetti, *J. Med. Chem.*, 2019, **62**, 2218–2244.

45 A. Geerlof, *M9 mineral medium*, **https://www.helmholtz-muenchen.de/fileadmin/PEPF/Protocols/M9-medium_150510.pdf**.

46 W. Lee, M. Tonelli and J. L. Markley, *Bioinformatics*, 2015, **31**, 1325–1327.

47 M. P. Williamson, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2013, **73**, 1–16.

48 J. J. Helmus and C. P. Jaroniec, *J. Biomol. NMR*, 2013, **55**, 355–367.

49 A. Keeley, P. Ábrányi Balogh and G. M. Keserű, *Med. Chem. Commun.*, 2019, **10**, 263–267.

50 W. Kabsch, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 125–132.

51 A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658–674.

52 M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2011, **67**, 235–242.

53 P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 486–501.

54 D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams and P. D. Adams, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2019, **75**, 861–877.

55 A. Immirzi, *Crystallographic Computing Techniques*, ed. F. R. Ahmed, Munksgaard, 1966, p. 399.

56 L. F. Ten Eyck, *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.*, 1973, **29**, 183–191.

57 R. J. Read and A. J. Schierbeek, *J. Appl. Crystallogr.*, 1988, **21**, 490–495.

58 F. Long, R. A. Nicholls, P. Emsley, S. Gražulis, A. Merkys, A. Vaitkus and G. N. Murshudov, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2017, **73**, 112–122.

59 F. Long, R. A. Nicholls, P. Emsley, S. Gražulis, A. Merkys, A. Vaitkus and G. N. Murshudov, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2017, **73**, 103–111.
60 *TURBOMOLE V7.2 2017*, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007, available from **https://www.turbomole.com**.
61 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297.
62 C. Hättig, *Phys. Chem. Chem. Phys.*, 2005, **7**, 59–66.
63 M. Feyereisen, G. Fitzgerald and A. Komornicki, *Chem. Phys. Lett.*, 1993, **208**, 359–363.
64 F. Weigend, M. Häser, H. Patzelt and R. Ahlrichs, *Chem. Phys. Lett.*, 1998, **294**, 143–152.

