JBC RESEARCH ARTICLE

Enhanced identification of small molecules binding to hnRNPA1 *via* cryptic pockets mapping coupled with X-ray fragment screening

Received for publication, September 24, 2024, and in revised form, February 13, 2025 Published, Papers in Press, February 19, 2025,
https://doi.org/10.1016/j.jbc.2025.108335

Louise Dunnett[1], **Sayan Das**[2,3], **Vincenzo Venditti**[2,3], **and Filippo Prischi**[4,*]

From the [1]*Diamond Light Source Ltd., Harwell Science and Innovation Campus, Didcot, UK;* [2]*Department of Chemistry, Iowa State University, Ames, Iowa, United States;* [3]*Roy J. Carver Department of Biochemistry, Biophysics and Molecular Biology, Iowa State University, Ames, Iowa, United States; and* [4]*Randall Centre for Cell and Molecular Biophysics, King's College London, London, UK.*

Reviewed by members of the JBC Editorial Board. Edited by Karin Musier-Forsyth

The human heterogeneous nuclear ribonucleoprotein (hnRNP) A1 is a prototypical RNA-binding protein essential for regulating a wide range of post-transcriptional events in cells. As a multifunctional protein with a key role in RNA metabolism, deregulation of its functions has been linked to neurodegenerative diseases, tumor aggressiveness, and chemoresistance, which has fuelled efforts to develop novel therapeutics that modulate its RNA-binding activities. Here, using a combination of molecular dynamics simulations and graph neural network pocket predictions, we showed that hnRNPA1 N-terminal RNA-binding domain (unwinding protein 1 [UP1]) contains several cryptic pockets capable of binding small molecules. To identify chemical entities for the development of potent drug candidates and experimentally validate identified druggable hotspots, we carried out a large fragment screening on UP1 protein crystals. Our screen identified 36 hits that extensively sample UP1 functional regions involved in RNA recognition and binding as well as map hotspots onto novel protein interaction surfaces. We observed a wide range of ligand-induced conformational variation by stabilization of dynamic protein regions. Our high-resolution structures, the first of an hnRNP in complex with a fragment or small molecule, provide rapid routes for the rational development of a range of different inhibitors and chemical tools for studying molecular mechanisms of hnRNPA1-mediated splicing regulation.

RNA-binding proteins (RBPs) play a vital role in regulating post-transcriptional events in every cell. RBP recognition and binding to RNA sequences and/or secondary structure motifs are highly dynamic and tightly regulated processes. In humans, there are about 2000 RBPs, grouped into several families based on the presence of structurally conserved RNA-binding domains (1, 2). The heterogeneous nuclear ribonucleoprotein (hnRNP) A1 is a prototypical RBP implicated in multiple aspects of nucleic acid metabolism, including processing of miRNA precursors, transcription regulation, constitutive and alternative splicing, and nucleocytoplasmic mRNA transport (3). HnRNPA1 ability to carry out these wide range of functions is encoded in its modular domain organization consisting of an N-terminal unwinding protein 1 (UP1) domain and a C-terminal intrinsically disordered glycine-rich domain (4–6). UP1 is composed of two RNA recognition motifs (RRM1 and RRM2), which, despite having a high level of sequence similarity and recognizing similar optimal motifs (5′-YAG-3′, where Y is a pyrimidine), are nonredundant and functionally nonequivalent (7). Over the last 20 years, several structural studies of the protein in complex with telomeric DNA, RNA, and miRNA (6, 8, 9) have provided ligand binding models, although the molecular mechanisms driving selective DNA–RNA recognition remain elusive (10).

As a major regulator of gene expression, deregulation of hnRNPA1 functions has been linked to several pathological cellular conditions. A growing number of studies have suggested that hnRNPA1 contributes to development of neurodegenerative diseases (11). For example, hnRNPA1 mutations have been identified to alter splicing, often by causing exon-skipping events, and to increase the formation of self-seeding fibrils in multiple sclerosis (12) and amyotrophic lateral sclerosis (13), respectively. Furthermore, altered hnRNPA1 expression has been found to enhance translation of dynamin-related protein 1 (Drp1) (14), a GTPase protein often overexpressed in Huntington's disease, and affect the alternative splicing of the amyloid precursor protein, causing increased secretion of β-amyloid peptide (15). HnRNPA1 has also been shown to promote viral replication. Specifically, several viruses (*e.g.*, human rhinovirus, enterovirus, sindbis virus, and human immunodeficiency virus I) exploit hnRNPA1 internal transactivating factor activity, resulting in an increased internal ribosome entry site–mediated translation of viral RNA (16–19). Deregulated hnRNPA1 expression or activation *via* post-translational modifications has been shown to increase cell proliferation and survival in various cancer types, including lung, breast, prostate, and gastric cancers (20–24), leukemia (25), Burkitt lymphoma (26), multiple myeloma (27), and hepatocellular and cervical carcinomas (28, 29). Although

* For correspondence: Filippo Prischi, filippo.prischi@kcl.ac.uk.

hnRNPA1 has been shown to be involved in a wide range of molecular events driving tumorigenesis and drug resistance, the underlying molecular mechanisms seem to be linked to an alteration in RNA recognition and binding, leading to an increased translation of prosurvival proteins and oncoprotein variants (3). Consistently, it has been shown that knockdown of hnRNPA1 preferentially leads to apoptosis in cancer cells over their healthy counterparts (30).

Due to the direct and indirect involvement of hnRNPA1 in neurodegenerative disorders, viral gene expression, and carcinogenesis, preventing or fine-tuning hnRNPA1–RNA binding with novel therapeutics is an emerging area of scientific interest, especially considering the growing number of RBP inhibitors showing promising preclinical activities (31). Earlier attempts to target hnRNPA1 identified a small number of compounds that alter the protein activity in different ways. For example, Carabet *et al.* (32) showed that the small molecule VPC-80051 binds to the RRM1 of hnRNPA and may act as a RNA competitor, causing a small reduction of the androgen receptor splice variant (32). Quercetin, a promiscuous naturally occurring flavonoid, has been shown to bind the unstructured C-terminal region of hnRNPA1, preventing hnRNPA1 nucleocytoplasmic shuttling and causing cytoplasmic accumulation (33). Similarly, camptothecin, a noncompetitive protein–protein interaction (PPI) inhibitor, binds the hnRNPA1 C-terminal region and prevents its interaction with topoisomerase I (34). Although promising, these compounds show limited selectivity and specificity for hnRNPA1. A major obstacle in the identification of suitable inhibitors is the lack of easily identifiable deep binding pockets. Here, to overcome this limitation and identify small-molecule ligands of hnRNPA1, we used molecular dynamics (MD) simulations and graph neural network (NN) pocket predictions coupled with X-ray fragment screening. Our simulations showed that hnRNPA1 N-terminal RNA-binding domain (UP1) contains several cryptic pockets (*i.e.*, concavities, often absent in crystal structures, that open on the protein surface when the protein fluctuates to an excited state) capable of binding small molecules. Using the XChem facility at Diamond Light Source, we carried out a large fragment screening on UP1 protein crystals. Our screen identified 36 hits, which extensively samples UP1 functional regions involved in RNA recognition and binding as well as map hotspots (*i.e.*, a cluster of residues that makes a major contribution to the binding free energy) onto novel protein interaction surfaces. Our high-resolution structures, the first of an hnRNP in complex with a fragment or small molecule, provides rapid routes for the rational development of a range of different potent inhibitors and chemical tools for studying molecular mechanisms of hnRNPA1-mediated splicing regulation.

Results

Computational hotspot mapping

MD-derived structures have been extensively used to sample small conformational rearrangements and map hotspots and cryptic pockets on protein surfaces, showing good agreement between predicted and experimentally mapped pockets (35, 36). We focused our studies on the UP1 domain, as hnRNPA1 C-terminal region is intrinsically disordered and not suitable for structural studies. Differently, the two RRMs in UP1 have a conserved βαββαβ structure with a four-stranded antiparallel β-sheet sitting over two helices, joined together by a highly dynamic inter-RRM loop (Fig. 1*A*). The two RRMs are in close contact, with the α2 helix of one domain interacting with the α2 helix of the other domain in an antiparallel orientation (4–6).

To identify cryptic pockets on UP1 surface we performed a 2 μs MD simulation, which allows to observe how protein conformational dynamics in a water environment contribute to the formation of binding cavities. In line with previous similar studies (37), we observed good overall agreement between the root mean square fluctuation derived from our MD simulations and the root mean square fluctuation calculated using the experimentally determined UP1 NMR structural ensemble (Protein Data Bank [PDB] ID: 2LYV) (Fig. S1). Thus, we generated an ensemble of structures using the MD trajectories, which were used to map and characterize cryptic pockets using fpocket 2.0 (38). To further validate identified pockets, we compared fpocket results with pockets predicted using PocketMiner, which utilizes a graph NN trained using structure ensembles sampling pockets opening events (39). This robust approach increased reliability that the predicted pockets exist on UP1 and that these pockets are also suitable for small molecule binding.

Both approaches identified similar pockets located on UP1 functional regions. Specifically, two hotspots were identified on the two UP1-binding platforms or nucleobase pockets (6), created by the β-sheets of the RRM1 and the inter-RRM loop (RRM1 nucleobase pocket), and the β-sheets of the RRM2 and the UP1 C-terminal (RRM2 nucleobase pocket), highlighting the key roles that the highly dynamic inter-RRM loop and the UP1 C-terminal play in the formation of these pockets (Fig. 1, *B* and *C*). A small difference is visible for the RRM2, where our MD simulations identified a large cryptic pocket that extends across the whole RRM2 β-sheet, whereas PocketMiner predicts a smaller pocket, which does not include the region in between β1 and β3 (Fig. 1*C*).

The largest hotspot predicted by the MD simulation extends throughout the whole interface between the two RRMs (inter-RRM interface or inter-RRM [α2–α2]), whereas PocketMiner suggests that only part of this surface is viable for small-molecule binding (Fig. 1, *B* and *C*). This inter-RRM interface is a known allosteric site and is key for hnRNPA1 function. Indeed, mutation studies disrupting interactions stabilizing the inter-RRM interface (two salt-bridges between R88–D157 and R75–D155 and a hydrophobic cluster involving L13, I164, and V90) caused a reduction in RNA-binding affinity and overall protein stability (6, 10).

Finally, smaller binding pockets are visible at the edges of the two RRMs, likely generated by dynamics of the flexible RRMs β2–β3 loops (Fig. 1*B*), none of which were detected by PocketMiner (Fig. 1*C*).

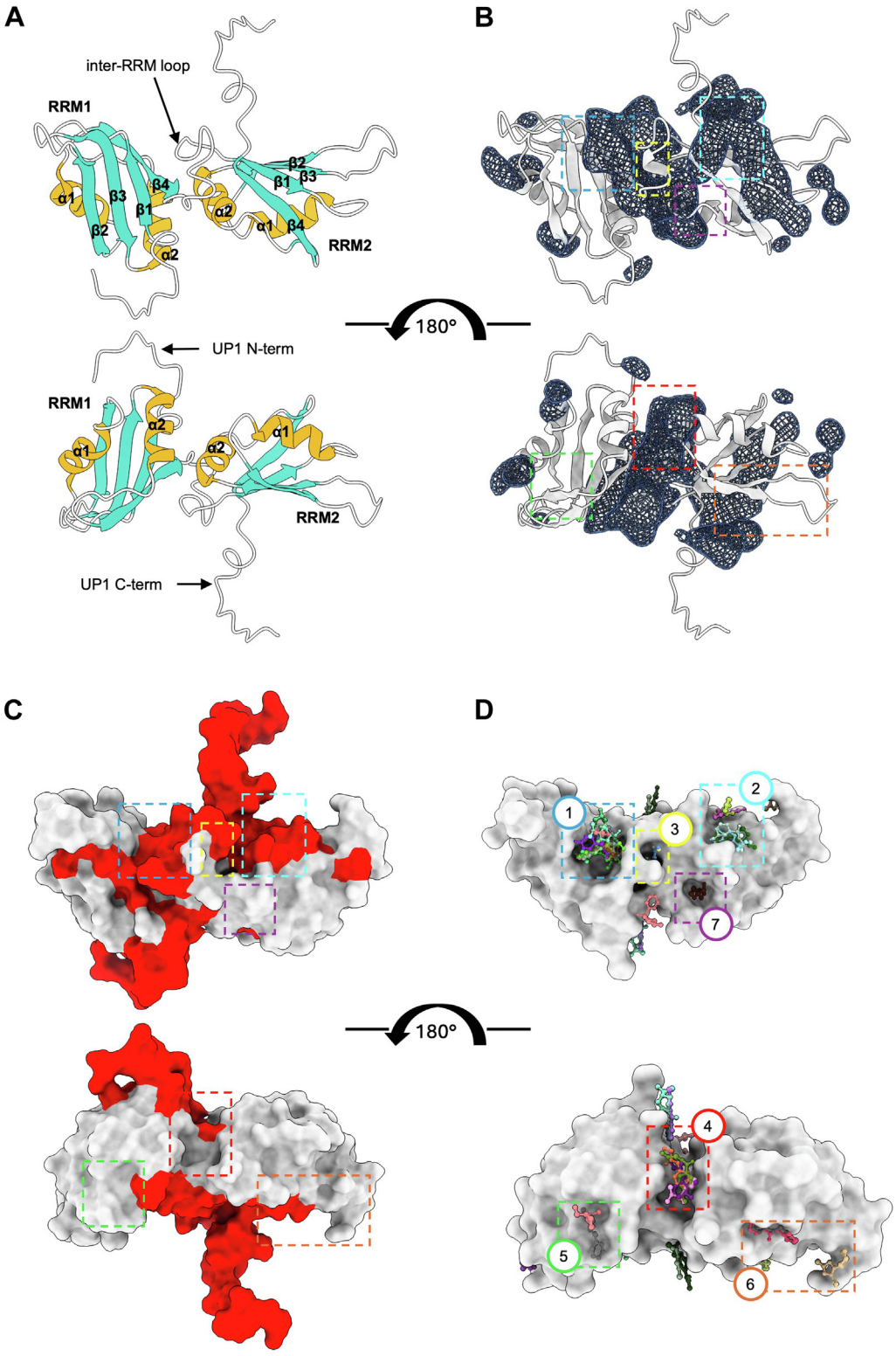

Figure 1. UP1 cryptic pocket identification. *A*, *cartoon representation* of UP1, with labeled conserved secondary structures. Flexible regions (N and C termini and the inter-RRM loop) are visible in the NMR structure but not in most X-ray structure of the apo UP1. *B*, UP1 *cartoon representation* (*gray*) with cryptic pockets identified *via* MD simulations, shown as *blue isosurfaces*. *C*, UP1 surface colored according to PocketMiner predicted cryptic pocket likelihood, >0.5 *red* and <0.5 *gray*. *D*, UP1 crystal structure with all fragment hits shown as *ball and sticks*, and druggable hotspots highlighted by colored boxes. For reference, colored boxes in (*D*) are also shown on (*B*) and (*C*). **1**: nucleobase pocket 1; **2**: nucleobase pocket 2; **3**: inter-RRM loop (hidden behind the loop in (*A*) and (*B*)); **4**: inter-RRMs (α2–α2); **5**: RRM1 α1–(α2–β4) loop; **6**: RRM2 α1–β2 and the adjacent RRM2 (β2–β3) loop sites; **7**: RRM2 α2–β4. All structures shown have the same orientation. MD, molecular dynamics; UP1, unwinding protein 1.

Taken together, these data suggest that several druggable cryptic pockets exist on the hnRNPA1 N-terminal RNA-binding domain surface, which encouraged us to carry out a crystallographic fragment screening.

Large fragment screening on UP1 protein crystals

There is currently no experimental structural information for hnRNPA1, or any other member of the hnRNP family, in complex with small molecules or drug-like compounds. To overcome this limitation, we carried out high-throughput fragment screening on protein crystals to identify chemical entities for the development of potent drug candidates and to chemically probe protein surface for the identification of druggable hotspots (40, 41). We screened the Diamond-SGC-iNext Poised Library, comprising 768 chemically diverse fragments (42, 43), directly on UP1 crystals. Of 548 successfully collected datasets, PanDDA (44) identified 61 potential hits, which were manually verified in Coot (45), resulting in 36 refined high-resolution UP1 fragment–bound structures (Table S1). Nearly 80% of all the identified hits occupy known functional regions of UP1 and computationally predicted binding sites (Fig. 1*D*), with 48% and 13% of identified fragments binding to the RRM1 and two nucleobase pockets, respectively, 13% on the inter-RRM (α2–α2) and 6% on the inter-RRM loop (Fig. 1*D*). The high level of agreement between experimentally identified hotspots and predicted cryptic pockets increases reliability of identified hits and released models and improves accuracy in the prioritization of hits. The remaining hits are found onto a likely conserved RRM PPI surface (Table S3). The majority of compounds showed high binding surface selectivity, and only seven fragments bind at the interface of two or three symmetry-related molecules, two of which were also identified on two different binding surfaces (Table S3).

Ligand-induced conformational variations on the RRM1

Fifteen fragments were identified to bind to the RRM1 nucleobase pocket, all interacting with key residues in the two ribonucleoprotein consensus sequences (RNPs). Indeed, extensive structural studies investigating UP1 mode of binding to the consensus motifs 5′-YAGG-3′ have shown that RNPs mediate sequence-specific ssRNA binding. Specifically, the conserved phenylalanines in the RNPs (F17, F57, and F59) directly interact with the central A_2G_3 nucleotides, with the central A_2 stacking between F17 and H101, the G_3 stacking onto F59 and forming van der Waals contacts with F57, and the phosphate group between A_2 and G_3 forming an electrostatic interaction with R55 (8, 46). Additional residues in the RRM1 nucleobase pocket, E85/K87 and D42/R92, form a network of hydrogen bonds with the flanking Y_1 and G_4, respectively (6, 8, 46). Interestingly, all fragments identified on the RRM1 nucleobase pocket formed van der Waals interactions with the conserved F17, F57, and/or F59 and frequently π-stacking with F17 (Table S4) through aromatic rings or similar heterocycles. Importantly, we observed a series of highly relevant conformational variations in the side chains of the key conserved F17 and F59 in the RNPs of RRM1 (Table S3). Earlier studies by Vitali *et al.* (47) showed that in the UP1 apo structure, F17 and F59 have correlated alternative conformations, with side chains adopting (i) F17A–F59A orientation (RNA-bound conformation), seen also in the UP1-DNA and -RNA bound structures and the only conformation seen in other RRM–nucleic acid complex structures (6, 8) and (ii) F17B–F59B (anti–RNA-bound conformation), not frequently seen in other RRM structures, including the RRM2 of UP1. Fragments binding to the RRM1 nucleobase pocket modulate these dynamic features by recognizing different RRM1 RNP conformations (Table S3). Three fragments (Z1373445602, Z1401276297, and Z106579662) sit on the RRM1 β-sheet in an orientation that closely resemble that of the purine ring of the central A_2, with their aromatic rings sandwiched between F17 and H101, stabilizing the F17 side chain in an RNA-bound conformation but not altering F59 dynamics (Fig. 2*A*). We also observed the opposite conformational change, likely induced by the presence in the fragments (ZINC72259689, Z137811222, and Z1217960891) of a large cycle at one end of the fragment positioned in proximity of the F59, which locks its side chain in an anti–RNA-bound conformation, and a nitrogen containing cycle at the other end, which interacts with either conformation of the F17 (Fig. 2*B*). Three fragments are positioned parallel to the RRM1 β-sheet direction, in an orientation similar to that occupied by the sugar–phosphate backbone in the UP1-DNA-bound structures (6, 8). These three fragments (Z56880342, Z54508609, and Z1220452176) have a similar geometry and group composition, with an aromatic head stacked in the middle of F17, F57, and F59 and a polar tail protruding outward. This induces a conformational variation of the side chains of F17 and F59, which adopt (mainly or only) an anti–RNA-bound conformation (Fig. 2*C*) (6, 8). Only four fragments reach the inter-RRM loop and alter the RRM1 nucleobase pocket structure. (i) ZINC72259689 and Z1203107138 form H-bonds with the V90 backbone, partially rigidifying the inter-RRM loop and pushing the H101 side chains backward (toward the RRM2) of 7.5 Å compared with the Cγ relative position in the DNA-bound structure (Fig. 2*B*) (residue not visible in the apo structure). Differently (ii) Z991506900 and Z30820160 coordinate several water molecules, with a much higher number of visible water molecules (not present in other structures) bridging the inter-RRM loop. This is likely linked to a reduction in flexibility of the loop (fully visible in both structures), which folds back onto the RRM1 nucleobase pocket, creating a pronounced positive charged binding cavity in which the fragments are bound (Figs. 2*D* and S1*A*). Interestingly, this also induces an RNA-bound conformation for both F17 and F59 side chains (Table S3).

Fragment hits reveal multiple targetable hotspots

Similar to RRM1, fragments binding to the RRM2 nucleobase pocket interact with the hydrophobic residues on the RNPs (F108 and F148) (Fig. 3*A*), occupying a position similar to that of purine rings of either the A_2 (Z416341642,

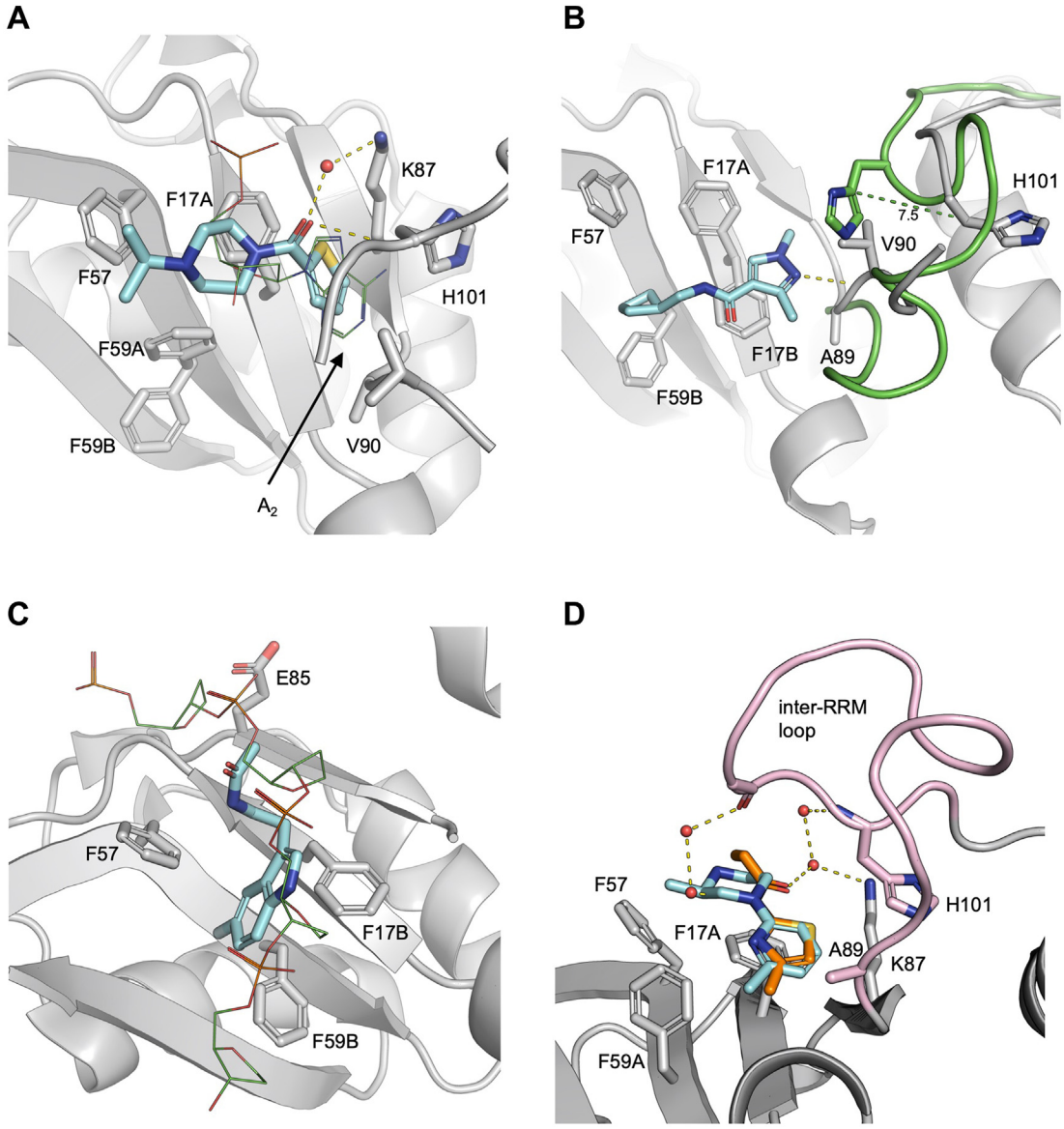

Figure 2. RRM1 nucleobase pocket. *A*, overlay of Z106579662 (*cyan sticks*) and A_2 (PDB ID: 2UP1) (*green lines*). Residues mediating interaction are shown as *gray sticks*, highlighting F59 alternative conformations and F17 RNA-bound conformation. *B*, ZINC72259689 (*cyan sticks*) stabilizes F59 in anti–RNA-bound conformation but does not affect F17 alternative conformations. The UP1-ZINC72259689 structure (*gray cartoon*) is overlayed to UP1-DNA structure (PDB ID: 2UP1, with inter-RRM loop shown as *green cartoon*). The *dotted green line* highlights the different position of H101 side chain in the two structures. *C*, overlay of Z1220452176 (*cyan sticks*) and the DNA sugar-phosphate backbone (PDB ID: 2UP1) (*green lines*), highlighting F17 and F59 anti–RNA-bound conformations. *D*, overlay of Z30820160 (*orange sticks*) and Z991506900 (*cyan sticks*), with F17 and F59 in RNA-bound conformations. The fully visible inter-RRM loop is shown in *pink*. In all panels, water molecules are shown as *red spheres*, *sticks* are colored by heteroatom, and hydrogen bonds with lengths 2.5–3.5Å are shown as *yellow dotted lines*. Electron density maps of ligands are shown in Fig. S6. PDB, Protein Data Bank; UP1, unwinding protein 1.

Z992569480, and Z906021418) or G_3 (Z641230552 and Z641239276).

Differently, binding of fragments on the inter-RRM loop and the inter-RRM (α2–α2) is mostly driven by polar interactions. Two fragments (Z57040482 and Z237527902) bind directly to the inter-RRM loop. This induces a conformational variation causing the loop to fold toward RRM2 α2–β4 and form a deep pocket (Fig.S2*B*), similar to the one present in the UP1-RNA/DNA-bound structures (6, 8) but absent in the apo structure (4, 5). These fragments form H-bonds between the backbone amino group of L102 backbone, the side chain of K179, and a carbonyl group in an acetamide moiety, and between the backbone amino group of R92 and the side chain of S91 and the phenyl group in Z57040482 and the nitrogen-containing heterocycle of Z237527902 (Fig. 3*B*, Table S4). The four fragments identified in the inter-RRM (α2–α2) contain one or two amide or sulfonamide groups aligned parallel to the inter-RRM interface and forming a series of hydrogen bonds or water bridges with several residues from both the RRM1 (*i.e.*, D69, N73, R75) and the RRM2 (*i.e.*, E126, Q127, K161) (Fig. 3*C*, Table S4). These hits have expansion vectors suitable for targeting the allosteric pocket containing the inter-RRM salt-bridges (Fig. S3).

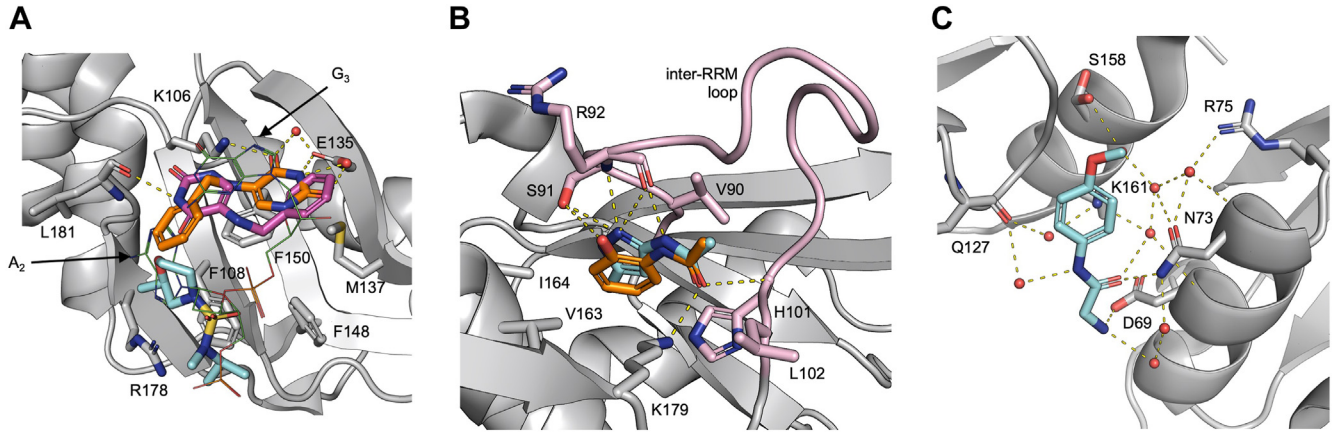

Figure 3. Fragments extensively sample UP1 functional regions. *A*, overlay of Z416341642 (*cyan sticks*), Z641230552 (with the two alternative locations in *orange* and *purple*), and DNA (PDB ID: 2UP1) (*green lines*) bound to the nucleobase pocket 2. *B*, overlay of Z57040482 (*orange sticks*) and Z237527902 (*cyan sticks*) bound to the inter-RRM loop site. The fully visible inter-RRM loop is shown in *pink*. *C*, inter-RRM (α2–α2) site with EN300-805013 (*cyan sticks*). In all panels, *sticks* are colored by heteroatom, water molecules are shown as *red spheres*, and hydrogen bonds with lengths 2.5–3.5Å are shown as *yellow dotted lines*. Electron density maps of ligands are shown in Fig. S6. UP1, unwinding protein 1.

RRM PPI surface binders

Earlier studies have shown that hnRNPA1 interacts with several macromolecules, yet there are no experimental structural data on hnRNPA1 in complex with proteins, peptides, or small molecules. However, structural studies on other RRM proteins (*i.e.*, FIR, PTBP1, hnRNPL, RBM7, CBP20) (48–53) have shown that the surface defined by α1–α2–(α2–β4) loop is a PPI surface. In all structures, a conserved hydrophobic depression on the PPI surface drives protein–peptide binding *via* hydrophobic interactions. This hydrophobic depression is also present on both RRM domains of UP1 (Fig. S4). Considering the high level of conservation of the RRM fold, it is tempting to speculate that these regions are also key in mediating PPI in UP1. While PPI surfaces lack the classic small molecules binding site features detected by MD hotspot mapping and PocketMiner, PPI surfaces can be targeted with small molecules that inhibit/interfere with PPI (54, 55). Importantly, this would recapitulate the presence of five fragments on the RRM1 (Z45617795 and EN300-115958) and RRM2 (Z86417414, Z802821712, and Z734147462) α1–α2–(α2–β4) loop PPI surfaces. (i) Z45617795 and EN300-115958 contain a methanesulfonamide group that forms H-bond with K78, and the phenyl group of Z45617795 is further stabilized by van der Waals interactions with P76 and V83 (Fig. 4*A*). (ii) The amide group of Z802821712 forms H-bonds with E118 and I134, whereas R22 forms a H-bond with nitrogen atoms in the pyridazine ring and a halogen bond with the fluorobenzene ring of Z86417414 and Z802821712, respectively (Fig. 4*B*), thus providing an accessible growth vector for targeting the RRM2 PPI surface. Z734147462 forms H-bonds between the oxygen atoms of the tetrahydropyran–methanol group and Q165 and V177 (Fig. 4*C*). Interestingly, this PPI surface is present also in noncanonical RRM domains, which, for example, has been exploited for the design of cyclic peptide and small-molecule splicing inhibitors targeting the splicing factor SPF45 (56, 57).

Fragment interactions influenced by crystal contacts

Søndergaard *et al.* (58) showed that about a third of ligand-bound structures in the PDB have ligands positioned at less

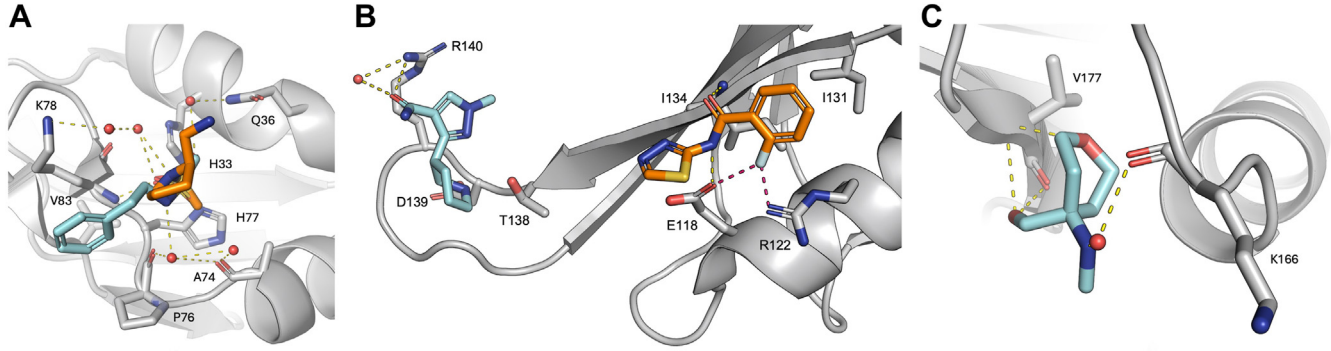

Figure 4. RRM protein–protein interaction (PPI) surface binders. *A*, overlay of Z45617795 (*cyan sticks*) and EN300-115958 (*orange sticks*) bound to the RRM1 α1–(α2–β4) loop site. *B*, RRM2 α1–β2 site with Z86417414 (*orange sticks*) and the adjacent RRM2 (β2–β3) loop site with EN300-197154 (*cyan sticks*). Halogen bonds (2.3–2.8Å) are visualized as *pink dotted lines*. *C*, RRM2 α2–β4 site with Z734147462 (*cyan sticks*). In all panels, *sticks* are colored by heteroatom, water molecules are shown as *red spheres*, and hydrogen bonds with lengths 2.5–3.5Å are shown as *yellow dotted lines*. Electron density maps of ligands are shown in Fig. S6.

than 5 Å distance to a symmetry-related atom, and their binding is influenced by crystal contacts and packing. We performed a stringent assessment of the models and identified seven fragments stabilized by interactions with residues from symmetry-related molecules. Of these, six fragments (Z1491353358, Z235361315, EN300-118084, Z198195770, Z104584152, and Z106579662) are primarily bound to the inter-RRM (α2–α2) and in addition stabilized by interactions with residues in the (β2–β3) loop (Fig. S5, *A* and *B*), whereas one (Z641239276) is bound to the RRM2 nucleobase pocket and is in addition stabilized by interactions with residues in the RRM2 α1–β3 (Fig. S5*C*). Of particular interest is Z1491353358, which allosterically induces conformational variations in the RRM1 RNPs (Fig. S2*A*). Specifically, Z1491353358 induces an RNA-bound conformation for the F17 and F59 side chains, potentially as a result of its close proximity to the inter-RRM loop, which shifts (1.6 Å and 2.5 Å compared with the A89 relative position in UP1 apo and DNA-bound structure, respectively) toward the RRM1 RNPs. This results in an increased molecular crowding, which may force the F17 and F59 side chains into the RNA-bound conformation.

Discussions

Despite advances in protein–RNA inhibition *via* RNA targeting (*i.e.*, antisense oligonucleotides and short interfering RNAs), RBP small-molecule inhibitors are scarce, mostly because of the lack of easily identifiable pockets suitable for binding drug-like molecules (59). Most small-molecule inhibitors targeting RBPs have been identified through standard high-throughput screening of large compound libraries (59). However, an earlier study on hepatitis C virus NS5b RNA-dependent RNA polymerase showed that fragments with millimolar affinities, identified using X-ray fragment screening, could be developed into RBP inhibitors with nanomolar affinities using structure-based design (60). No similar studies have been carried out on other RBPs, and only a handful of similar X-ray screenings have been carried out on DNA-binding proteins (61, 62). In this article, we performed a crystallography-based fragment screen study on a prototypical RBP, the N-terminal RNA-binding domain of hnRNPA1 (UP1). We first analyzed UP1 protein surface using a combination of MD simulations and graph NN predictions, and our simulations highlighted the presence of several different cryptic pockets with ideal features for small-molecule interactions. To identify compounds able to bind these pockets, we carried out a high-throughput fragment screening on protein crystals, and we solved the first experimental structures of hnRNPA1 (also the first structure of a member of the hnRNP family and an RRM) in complex with small molecules. The bound fragments interact with key protein surfaces, including RNA-binding sites, allosteric regions, and PPI surfaces.

It has been shown that proteins are able to assume holo-like conformations even in the absence of interacting ligands (36). Thus, cryptic pockets are present in the apo structures of proteins and become accessible because of protein dynamics. The ligand then recognizes its target by "selecting" the most complementary conformation from an ensemble of metastable states and/or by "induced fit," in either case causing a population shift toward holo states (36). A growing number of studies have shown that these holo-like states can be identified using MD simulations (63, 64). Here, we performed long MD simulations (2 μs) to identify UP1 holo-like states. To robustly and reliably identify cryptic pockets on UP1, we compared MD simulation results with pockets predicted using PocketMiner (39). The high level of agreement between predicted pockets strongly suggests that there are several druggable sites on UP1 surface and encouraged us to carry out an X-ray fragment screening. Indeed, the current lack of experimental structural information is impairing the design and/or identification of compounds with high selectivity or affinity for hnRNPA1.

Small molecules are the most convenient route to target RBPs, as they present advantageous features, including good solubility and oral bioavailability, ability to cross cell membranes and reach the different compartments where RBPs are localized (65). Despite the lack of experimental structural information, several small-molecule inhibitors able to bind to RRM domains and modulate RBP functions have been developed (2). These include compound 27 (and derivatives 838 and 841), which binds directly to the RRM2 of UGBP Elav-like family member 1 (CELF1) and competes with interferon-γ mRNA (66). This prevents mRNA degradation and inhibits hepatic stellate cell activation, suggesting it could be a viable treatment for liver fibrosis (66). Earlier studies identified MS-444 as an inhibitor of Hu antigen R (HuR) dimerization, probably *via* binding to its RRM1/2 domains, thus interfering with RNA binding and HuR cellular localization (67). With a slightly different mechanism, compound 1c (and derivatives) was predicted to bind in between RRM1 and RRM2 of HuR and compete with RNA binding, showing to potentiate the antitumor effect of standard chemotherapy treatment in breast cancer xenografts (68). Similarly, the small molecule Ro binds to the RRM1 of MSI2 and compete with RNA binding, reducing survival of human acute myeloid leukemia cell lines without affecting normal cells (69). Ro was shown to be selective and not bind other RBPs (69). Indeed, the similarity among all published RRM inhibitors is low (Tanimoto coefficient <0.25—data not shown) (70), which would recapitulate the observed selectivity. Interestingly, also the similarity between available RRM inhibitors and our identified fragments is low (Tanimoto coefficient <0.25), which would suggest that compounds evolved from fragments described here are likely to be highly selective for hnRNPA1. Similarly to these RRM inhibitors, 61% of fragments we identified binds directly on the RNA-binding surfaces of UP1, suggesting these are likely valuable building blocks for the development of RNA competitors. Furthermore, the ligand-induced conformational variations on the RRM1 RNPs we identified may provide different avenues to probe RNA binding. Indeed, recent studies on the RRM domain of CUG-BP2 have shown that RNP aromatic side-chain motions are critical for RNA recognition (71). The authors also highlighted that these local dynamics are influenced by the conformation of neighbor

residues and the binding of RNA is essential to stabilize one uniform set of conformations (71). This is consistent with what we have observed in our structures, where the binding of fragments on the RRM1 nucleobase pocket directly locked F17 and F59 in different sets of conformations. More studies are warranted to better understand how these different conformations can be exploited for targeting small molecules. While several studies have successfully identified direct RNA-binding competitors, only a handful of allosteric RBP inhibitors has been developed. Clingman *et al.* (72) showed that the binding of ω-9 fatty acids in between the α1–α2 of the RRM of MSI1 induces a conformational change that allosterically inhibits RNA binding. Here, we showed that 21% of the identified fragments binds UP1 on regulatory regions outside the RNA-binding sites. Of particular interest are fragments binding onto the inter-RRM interface, as recent studies have shown that this surface is key in relaying thermodynamic stability across UP1 surface, and mutations disrupting RRM1–RRM2 coupling reduce RNA binding and likely hnRNPA1 ability to form dimers (10). Our MD simulations, in agreement with previous NMR structural studies, identified a series of key contacts among D69, N73, and R75, and K161, D155, H156, and D157 (10). Compounds Z235361315 and Z33546965 directly interact with these residues and prevent formation of inter- and intra-RRM interactions among N73, D157, and K161, providing potential routes for the development of allosteric inhibitors.

Very few articles discuss ligand-binding sites located on symmetry-related crystal-packing interfaces, as these are discarded as artifacts and often escapes analysis likely because of (i) not being biological relevant binding sites, potentially the result of high concentration of ligands in the soaking experiment (73) or (ii) having detrimental effects by disrupting contacts between symmetry-related molecules. The few exceptions to this are ligand binding on crystallographic oligomer interfaces (74) and molecular glue (75). While the seven ligands we identified on symmetry-related crystal-packing interfaces are unlikely to be viable starting point for further development, they bind in very close proximity to other fragments in the inter-RRM interface, thus could provide information for chemical merging or linking.

The ability to modulate hnRNPA1 activity in cells of most compounds was assessed (data not shown), but, likely because of the low millimolar binding affinity and the low solubility, none showed biological activity, in line with what reported in other similar fragment screening studies (60, 76). However, the fragments we identified sample extensively the protein surface, identifying novel hotspots and revealing diverse expansion vectors (Fig. S3). Overall, our structures provide many clear routes to developing a range of different potent inhibitors and chemical tool for studying molecular mechanisms of hnRNPA1-mediated splicing regulation.

Experimental procedures

MD simulations and hotspot mapping

A 2 μs simulation was performed by using Gromacs 2023 package (77) with Amber03 force field (78). The NMR solution structure (PDB ID: 2LYV) (79) was used as a starting conformation. The system was placed in a TIP3P water box with a 1 nm distance between the boundary of the box and the system. One Cl^- counterion was added to neutralize the charge. Energy minimization was carried out using the steepest descent method with a step size of 0.01 nm. Minimization was run until the maximum force fell below 1000 kJ/mol/nm for 50,000 steps. The system was equilibrated at 300 K for 100 ps with a timestep of 2 fs for 50,000 steps and at a constant pressure (1 atm) for 100 ps with a timestep of 2 fs for 50,000 steps, both using the leap frog integrator. LINCS (LINear Constraint Solver) was used to constrain the covalent hydrogen bonds with a Verlet cutoff scheme with a 1.0 nm cutoff radius for neighbor list. The particle mesh Ewald scheme was used to treat long-range interactions with a Fourier grid spacing of 0.16 nm. The modified Berendsen thermostat (velocity-rescale thermostat) and the Parrinello–Rahman barostat were used to maintain the temperatures at 300 K and 1 atm. To generate an ensemble of structures, the MD trajectory was clustered by k-means clustering using MDAnalysis (80). Fifty clusters were generated, and the representative structure of each cluster is available at https://github.com/fprischi/UP1_clusters. Cryptic pockets were identified by analyzing the MD trajectory using fpocket 2.0 (38).

The NMR ensemble (PDB ID: 2LYV) was analyzed using PocketMiner (39) and we consider only pockets with a likelihood (averaged for the ensemble) greater than 0.5.

UP1 preparation for structural studies

The N-terminal $(His)_6$UP1 fusion was expressed and purified as previously described (79). In short, BL21(DE3) (Invitrogen) were transformed with a pETM-14 vector containing the DNA sequence encoding the UP1 domain of hnRNP A1 (residues 2–196) (UniProt entry: P09651). Cells were grown in LB media overnight at 18 °C, resuspended in lysis buffer (50 mM Tris–HCl, pH 8.00, 300 mM NaCl, 0.01% [v/v] Triton X-100, 20 mM imidazole, 5% [v/v] glycerol, 5 mg/ml lysozyme from hen's egg white, 1.0 mg/ml DNase from bovine pancreas, EDTA-free protease inhibitor [Mini Tablets; Pierce]) and lysed by sonication. Cell lysate was centrifuged at 18,000g for 1 h at 4 °C, supernatant was loaded on a Ni–NTA column (Cytiva), and $(His)_6$UP1 was eluted with an imidazole gradient. The $(His)_6$ tag was removed *via* digestion overnight with Human Rhinovirus 3C protease (PreScission protease) at 4 units/ml at 4 °C. UP1 was then loaded on a HiTrap HP SP column (Cytiva) equilibrated with 20 mM Mes (pH 6.00), 5 mM β-mercaptoethanol, 1 mM EDTA, 5% (v/v) glycerol, and was eluted with an NaCl gradient. Fractions containing UP1 were pooled and loaded onto a HiLoad 16/60 Superdex 75 (Cytiva) equilibrated with crystallization buffer (20 mM Mes [pH 6.00], 150 mM NaCl, 5 mM β-mercaptoethanol, 1 mM EDTA, and 5% [v/v] glycerol).

UP1 sitting drop crystallization

Protein from a single purification batch was used for in crystallo fragment screening in 96-well MRC 2 lens

crystallization trays (SWISS-Sci). All reagents were pre-equilibrated to 16 °C prior to use. A liquid handling robot (Mosquito; TTP Labtech) was used to dispense UP1 (from a 15 mg/ml stock) and precipitant solution (0.1 M Tris–HCl pH 8.50, 25% [w/v] PEG-4000, and 8% [v/v] MPD (RS)-2-methyl-2,4-pentanediole) as 400 nl sitting droplets at a 1:3 ratio of protein:precipitant. Each plate was immediately incubated at 16 °C following protein dispensing.

Droplet targeting using TexRank and fragment soaking

Soaking, mounting, and data collections were performed following the XChem facility workflow (81). Crystal droplets were imaged using a Rock Imager (Formulatrix) and targeted for fragment droplet dispensing using TexRank software (82). A total of 768 fragments from the DSI poised library (42) (500 mM stock concentration in dimethyl sulfoxide) were dispensed into the crystal droplets using ultrasonic sound waves by a liquid handling robot (ECHO 550; Labcyte, Inc) (83), with a 100 mM final concentration of fragment (20% [v/v] final concentration of dimethyl sulfoxide) in the droplets. Crystals were incubated at 16 °C for 3 h before being mounted with the aid of a Crystal Shifter (Oxford Lab Technologies) and flash frozen in liquid nitrogen. Automated data collection was performed on the I04-1 beamline with a Pilatus 6M-F detector (Dectris) using a fixed wavelength of 0.92 Å and a beam size of 60 × 50 mm on cryo-cooled crystals.

Structure solution and refinement

The diffraction images were processed using *xia2 3dii* or *xia2 dials* (84, 85), followed by AIMLESS (86). UP1 crystals belonged to space group P 1 21 1 and consisted of one protein per asymmetric unit. After molecular replacement and refinement of the initial model, the resulting maps were analyzed by PanDDA (44) followed by model building using Coot 0.9.8.93 (45). The structures of confirmed hits were then refined using PHENIX 1.20.1 (87).

Data availability

Atomic coordinates and structure files for the UP1-fragment crystal structures have been deposited in the PDB (http://www.pdb.org/). See Table S1 for access codes.

Supporting information—This article contains supporting information.

Acknowledgments—F.P. acknowledge support from BBSRC grant (grant no.: BB/X018997/1) and The Leverhulme Trust grant (grant no.: RPG-2018-230). This work was supported by funds from the National Institute of General Medical Sciences with grant no. R35GM133488 (to V.V.).

Author contributions—F.P. conceptualization; V.V. and F.P. validation; L.D. and S.D. investigation; V.V. and F.P. writing–original draft; V.V. and F.P. writing–review & editing; V.V. and F.P. supervision; F.P. visualization; F.P. project administration; F.P. funding acquisition.

Funding and additional information—Data acquisition was carried out as part of the XCHEM project LB18944-1. We thank the Diamond XCHEM team for their support and expertise.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: hnRNP, heterogeneous nuclear ribonucleoprotein; HuR, Hu antigen R; MD, molecular dynamics; NN, neural network; PDB, Protein Data Bank; PPI, protein–protein interaction; RBP, RNA-binding protein; RRM, RNA recognition motif; UP1, unwinding protein 1.

References

1. Hentze, M. W., Castello, A., Schwarzl, T., and Preiss, T. (2018) A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.* **19**, 327–341
2. Li, Q., and Kang, C. (2023) Targeting RNA-binding proteins with small molecules: perspectives, pitfalls and bifunctional molecules. *FEBS Lett.* **597**, 2031–2047
3. Roy, R., Huang, Y., Seckl, M. J., and Pardo, O. E. (2017) Emerging roles of hnRNPA1 in modulating malignant transformation. *Wiley Interdiscip. Rev. RNA* **8**, e1431
4. Shamoo, Y., Krueger, U., Rice, L. M., Williams, K. R., and Steitz, T. A. (1997) Crystal structure of the two RNA binding domains of human hnRNP A1 at 1.75 A resolution. *Nat. Struct. Biol.* **4**, 215–222
5. Xu, R. M., Jokhan, L., Cheng, X., Mayeda, A., and Krainer, A. R. (1997) Crystal structure of human UP1, the domain of hnRNP A1 that contains two RNA-recognition motifs. *Structure* **5**, 559–570
6. Morgan, C. E., Meagher, J. L., Levengood, J. D., Delproposto, J., Rollins, C., Stuckey, J. A., *et al.* (2015) The first crystal structure of the UP1 domain of hnRNP A1 bound to RNA reveals a new look for an old RNA binding protein. *J. Mol. Biol.* **427**, 3241–3257
7. Jain, N., Lin, H. C., Morgan, C. E., Harris, M. E., and Tolbert, B. S. (2017) Rules of RNA specificity of hnRNP A1 revealed by global and quantitative analysis of its affinity distribution. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 2206–2211
8. Ding, J., Hayashi, M. K., Zhang, Y., Manche, L., Krainer, A. R., and Xu, R. M. (1999) Crystal structure of the two-RRM domain of hnRNP A1 (UP1) complexed with single-stranded telomeric DNA. *Genes Dev.* **13**, 1102–1115
9. Kooshapur, H., Choudhury, N. R., Simon, B., Muhlbauer, M., Jussupow, A., Fernandez, N., *et al.* (2018) Structural basis for terminal loop recognition and stimulation of pri-miRNA-18a processing by hnRNP A1. *Nat. Commun.* **9**, 2479
10. Levengood, J. D., Potoyan, D., Penumutchu, S., Kumar, A., Zhou, Q., Wang, Y., *et al.* (2024) Thermodynamic coupling of the tandem RRM domains of hnRNP A1 underlie its pleiotropic RNA binding functions. *Sci. Adv.* **10**, eadk6580
11. Clarke, J. P., Thibault, P. A., Salapa, H. E., and Levin, M. C. (2021) A comprehensive analysis of the role of hnRNP A1 function and dysfunction in the pathogenesis of neurodegenerative disease. *Front. Mol. Biosci.* **8**, 659610
12. Salapa, H. E., Thibault, P. A., Libner, C. D., Ding, Y., Clarke, J. W. E., Denomy, C., *et al.* (2024) hnRNP A1 dysfunction alters RNA splicing and drives neurodegeneration in multiple sclerosis (MS). *Nat. Commun.* **15**, 356
13. Kim, H. J., Kim, N. C., Wang, Y. D., Scarborough, E. A., Moore, J., Diaz, Z., *et al.* (2013) Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* **495**, 467–473
14. Park, S. J., Lee, H., Jo, D. S., Jo, Y. K., Shin, J. H., Kim, H. B., *et al.* (2015) Heterogeneous nuclear ribonucleoprotein A1 post-transcriptionally regulates Drp1 expression in neuroblastoma cells. *Biochim. Biophys. Acta* **1849**, 1423–1431

15. Donev, R., Newall, A., Thome, J., and Sheer, D. (2007) A role for SC35 and hnRNPA1 in the determination of amyloid precursor protein isoforms. *Mol. Psychiatry* **12**, 681–690
16. Lin, J. Y., Shih, S. R., Pan, M., Li, C., Lue, C. F., Stollar, V., *et al.* (2009) hnRNP A1 interacts with the 5' untranslated regions of enterovirus 71 and Sindbis virus RNA and is required for viral replication. *J. Virol.* **83**, 6106–6114
17. Levengood, J. D., Tolbert, M., Li, M. L., and Tolbert, B. S. (2013) High-affinity interaction of hnRNP A1 with conserved RNA structural elements is required for translation and replication of enterovirus 71. *RNA Biol.* **10**, 1136–1145
18. Cammas, A., Pileur, F., Bonnal, S., Lewis, S. M., Leveque, N., Holcik, M., *et al.* (2007) Cytoplasmic relocalization of heterogeneous nuclear ribonucleoprotein A1 controls translation initiation of specific mRNAs. *Mol. Biol. Cell* **18**, 5048–5059
19. Barrera, A., Ramos, H., Vera-Otarola, J., Fernandez-Garcia, L., Angulo, J., Olguin, V., *et al.* (2020) Post-translational modifications of hnRNP A1 differentially modulate retroviral IRES-mediated translation initiation. *Nucleic Acids Res.* **48**, 10479–10499
20. Pino, I., Pio, R., Toledo, G., Zabalegui, N., Vicent, S., Rey, N., *et al.* (2003) Altered patterns of expression of members of the heterogeneous nuclear ribonucleoprotein (hnRNP) family in lung cancer. *Lung Cancer* **41**, 131–143
21. Chen, Y. H., Liu, J., Wang, W., Xiang, L., Wang, J. D., Liu, S. D., *et al.* (2018) High expression of hnRNPA1 promotes cell invasion by inducing EMT in gastric cancer. *Oncol. Rep.* **39**, 1693–1701
22. Otsuka, K., Yamamoto, Y., and Ochiya, T. (2018) Regulatory role of resveratrol, a microRNA-controlling compound, in HNRNPA1 expression, which is associated with poor prognosis in breast cancer. *Oncotarget* **9**, 24718–24730
23. Nadiminty, N., Tummala, R., Liu, C., Lou, W., Evans, C. P., and Gao, A. C. (2015) NF-kappaB2/p52:c-Myc:hnRNPA1 pathway regulates expression of androgen receptor splice variants and enzalutamide sensitivity in prostate cancer. *Mol. Cancer Ther.* **14**, 1884–1895
24. Roy, R., Durie, D., Li, H., Liu, B. Q., Skehel, J. M., Mauri, F., *et al.* (2014) hnRNPA1 couples nuclear export and translation of specific mRNAs downstream of FGF-2/S6K2 signalling. *Nucleic Acids Res.* **42**, 12483–12497
25. Iervolino, A., Santilli, G., Trotta, R., Guerzoni, C., Cesi, V., Bergamaschi, A., *et al.* (2002) hnRNP A1 nucleocytoplasmic shuttling activity is required for normal myelopoiesis and BCR/ABL leukemogenesis. *Mol. Cell Biol.* **22**, 2255–2266
26. Brockstedt, E., Rickers, A., Kostka, S., Laubersheimer, A., Dorken, B., Wittmann-Liebold, B., *et al.* (1998) Identification of apoptosis-associated proteins in a human Burkitt lymphoma cell line. Cleavage of heterogeneous nuclear ribonucleoprotein A1 by caspase 3. *J. Biol. Chem.* **273**, 28057, 273, 33884-33884
27. Shi, Y. J., Frost, P. J., Hoang, B. Q., Benavides, A., Sharma, S., Gera, J. F., *et al.* (2008) IL-6-induced stimulation of c-myc translation in multiple myeloma cells is mediated by myc internal ribosome entry site function and the RNA-binding protein, hnRNP A1. *Cancer Res.* **68**, 10215–10222
28. Zhou, Z. J., Dai, Z., Zhou, S. L., Fu, X. T., Zhao, Y. M., Shi, Y. H., *et al.* (2013) Overexpression of HnRNP A1 promotes tumor invasion through regulating CD44v6 and indicates poor prognosis for hepatocellular carcinoma. *Int. J. Cancer* **132**, 1080–1089
29. Kim, Y. J., Kim, B. R., Ryu, J. S., Lee, G. O., Kim, H. R., Choi, K. H., *et al.* (2017) HNRNPA1, a splicing regulator, is an effective target protein for cervical cancer detection comparison with conventional tumor markers. *Int. J. Gynecol. Cancer* **27**, 326–331
30. Patry, C., Bouchard, L., Labrecque, P., Gendron, D., Lemieux, B., Toutant, J., *et al.* (2003) Small interfering RNA-mediated reduction in heterogeneous nuclear ribonucleoparticule A1/A2 proteins induces apoptosis in human cancer cells but not in normal mortal cell lines. *Cancer Res.* **63**, 7679–7688
31. Julio, A. R., and Backus, K. M. (2021) New approaches to target RNA binding proteins. *Curr. Opin. Chem. Biol.* **62**, 13–23
32. Carabet, L. A., Leblanc, E., Lallous, N., Morin, H., Ghaidi, F., Lee, J., *et al.* (2019) Computer-aided discovery of small molecules targeting the RNA splicing activity of hnRNP A1 in castration-resistant prostate cancer. *Molecules* **24**, 763
33. Ko, C. C., Chen, Y. J., Chen, C. T., Liu, Y. C., Cheng, F. C., Hsu, K. C., *et al.* (2014) Chemical proteomics identifies heterogeneous nuclear ribonucleoprotein (hnRNP) A1 as the molecular target of quercetin in its anti-cancer effects in PC-3 cells. *J. Biol. Chem.* **289**, 22078–22089
34. Manita, D., Toba, Y., Takakusagi, Y., Matsumoto, Y., Kusayanagi, T., Takakusagi, K., *et al.* (2011) Camptothecin (CPT) directly binds to human heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1) and inhibits the hnRNP A1/topoisomerase I interaction. *Bioorg. Med. Chem.* **19**, 7690–7697
35. Raich, L., Meier, K., Gunther, J., Christ, C. D., Noe, F., and Olsson, S. (2021) Discovery of a hidden transient state in all bromodomain families. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2017427118
36. Kuzmanic, A., Bowman, G. R., Juarez-Jimenez, J., Michel, J., and Gervasio, F. L. (2020) Investigating cryptic binding sites by molecular dynamics simulations. *Acc. Chem. Res.* **53**, 654–661
37. Reinknecht, C., Riga, A., Rivera, J., and Snyder, D. A. (2021) Patterns in protein flexibility: a comparison of NMR "ensembles", MD trajectories, and crystallographic B-factors. *Molecules* **26**, 1484
38. Le Guilloux, V., Schmidtke, P., and Tuffery, P. (2009) Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics* **10**, 168
39. Meller, A., Ward, M., Borowsky, J., Kshirsagar, M., Lotthammer, J. M., Oviedo, F., *et al.* (2023) Predicting locations of cryptic pockets from single protein structures using the PocketMiner graph neural network. *Nat. Commun.* **14**, 1177
40. Jacquemard, C., and Kellenberger, E. (2019) A bright future for fragment-based drug discovery: what does it hold? *Expert Opin. Drug Discov.* **14**, 413–416
41. Nichols, C., Ng, J., Keshu, A., Kelly, G., Conte, M. R., Marber, M. S., *et al.* (2020) Mining the PDB for tractable cases where X-ray crystallography combined with fragment screens can be used to systematically design protein-protein inhibitors: two test cases illustrated by IL1beta-IL1R and p38alpha-TAB1 complexes. *J. Med. Chem.* **63**, 7559–7568
42. Cox, O. B., Krojer, T., Collins, P., Monteiro, O., Talon, R., Bradley, A., *et al.* (2016) A poised fragment library enables rapid synthetic expansion yielding the first reported inhibitors of PHIP(2), an atypical bromodomain. *Chem. Sci.* **7**, 2322–2330
43. Douangamath, A., Powell, A., Fearon, D., Collins, P. M., Talon, R., Krojer, T., *et al.* (2021) Achieving efficient fragment screening at XChem facility at diamond light source. *J Vis Exp* **171**, e62414
44. Pearce, N. M., Krojer, T., and von Delft, F. (2017) Proper modelling of ligand binding requires an ensemble of bound and unbound states. *Acta Crystallogr. D Struct. Biol.* **73**, 256–266
45. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
46. Beusch, I., Barraud, P., Moursy, A., Clery, A., and Allain, F. H. (2017) Tandem hnRNP A1 RNA recognition motifs act in concert to repress the splicing of survival motor neuron exon 7. *Elife* **6**, e25736
47. Vitali, J., Ding, J., Jiang, J., Zhang, Y., Krainer, A. R., and Xu, R. M. (2002) Correlated alternative side chain conformations in the RNA-recognition motif of heterogeneous nuclear ribonucleoprotein A1. *Nucleic Acids Res.* **30**, 1531–1538
48. Cukier, C. D., Hollingworth, D., Martin, S. R., Kelly, G., Diaz-Moreno, I., and Ramos, A. (2010) Molecular basis of FIR-mediated c-myc transcriptional control. *Nat. Struct. Mol. Biol.* **17**, 1058–1064
49. Joshi, A., Coelho, M. B., Kotik-Kogan, O., Simpson, P. J., Matthews, S. J., Smith, C. W., *et al.* (2011) Crystallographic analysis of polypyrimidine tract-binding protein-Raver1 interactions involved in regulation of alternative splicing. *Structure* **19**, 1816–1825
50. Bhattacharya, S., Wang, S., Reddy, D., Shen, S., Zhang, Y., Zhang, N., *et al.* (2021) Structural basis of the interaction between SETD2 methyltransferase and hnRNP L paralogs for governing co-transcriptional splicing. *Nat. Commun.* **12**, 6452
51. Falk, S., Finogenova, K., Melko, M., Benda, C., Lykke-Andersen, S., Jensen, T. H., *et al.* (2016) Structure of the RBM7-ZCCHC8 core of the

NEXT complex reveals connections to splicing factors. *Nat. Commun.* **7**, 13573
52. Schulze, W. M., and Cusack, S. (2017) Structural basis for mutually exclusive co-transcriptional nuclear cap-binding complexes with either NELF-E or ARS2. *Nat. Commun.* **8**, 1302
53. Schmeing, S., Amrahova, G., Bigler, K., Chang, J. Y., Openy, J., Pal, S., *et al.* (2023) Rationally designed stapled peptides allosterically inhibit PTBP1-RNA-binding. *Chem. Sci.* **14**, 8269–8278
54. Lu, H., Zhou, Q., He, J., Jiang, Z., Peng, C., Tong, R., *et al.* (2020) Recent advances in the development of protein-protein interactions modulators: mechanisms and clinical trials. *Signal Transduct. Target. Ther.* **5**, 213
55. Ran, X., and Gestwicki, J. E. (2018) Inhibitors of protein-protein interactions (PPIs): an analysis of scaffold choices and buried surface area. *Curr. Opin. Chem. Biol.* **44**, 75–86
56. Jagtap, P. K., Garg, D., Kapp, T. G., Will, C. L., Demmer, O., Luhrmann, R., *et al.* (2016) Rational design of cyclic peptide inhibitors of U2AF homology motif (UHM) domains to modulate pre-mRNA splicing. *J. Med. Chem.* **59**, 10190–10197
57. Kobayashi, A., Clement, M. J., Craveur, P., El Hage, K., Salone, J. M., Bollot, G., *et al.* (2022) Identification of a small molecule splicing inhibitor targeting UHM domains. *FEBS J.* **289**, 682–698
58. Sondergaard, C. R., Garrett, A. E., Carstensen, T., Pollastri, G., and Nielsen, J. E. (2009) Structural artifacts in protein-ligand X-ray structures: implications for the development of docking scoring functions. *J. Med. Chem.* **52**, 5673–5684
59. Wu, P. (2020) Inhibition of RNA-binding proteins with small molecules. *Nat. Rev. Chem.* **4**, 441–458
60. Antonysamy, S. S., Aubol, B., Blaney, J., Browner, M. F., Giannetti, A. M., Harris, S. F., *et al.* (2008) Fragment-based discovery of hepatitis C virus NS5b RNA polymerase inhibitors. *Bioorg. Med. Chem. Lett.* **18**, 2990–2995
61. Wever, M. J. A., Scommegna, F. R., Egea-Rodriguez, S., Dehghani-Tafti, S., Brandao-Neto, J., Poisson, J. F., *et al.* (2024) Structure-based discovery of first inhibitors targeting the helicase activity of human PIF1. *Nucleic Acids Res.* **52**, 12616–12632
62. Yin, Z., Whittell, L. R., Wang, Y., Jergic, S., Liu, M., Harry, E. J., *et al.* (2014) Discovery of lead compounds targeting the bacterial sliding clamp using a fragment-based approach. *J. Med. Chem.* **57**, 2799–2806
63. Trampari, E., Prischi, F., Vargiu, A. V., Abi-Assaf, J., Bavro, V. N., and Webber, M. A. (2023) Functionally distinct mutations within AcrB underpin antibiotic resistance in different lifestyles. *NPJ Antimicrob. Resist.* **1**, 2
64. Bernini, A., Galderisi, S., Spiga, O., Bernardini, G., Niccolai, N., Manetti, F., *et al.* (2017) Toward a generalized computational workflow for exploiting transient pockets as new targets for small molecule stabilizers: application to the homogentisate 1,2-dioxygenase mutants at the base of rare disease Alkaptonuria. *Comput. Biol. Chem.* **70**, 133–141
65. Yang, N. J., and Hinner, M. J. (2015) Getting across the cell membrane: an overview for small molecules, peptides, and proteins. *Methods Mol. Biol.* **1266**, 29–53
66. Tan, Y., Sun, X., Xu, Y., Tang, B., Xu, S., Lu, D., *et al.* (2022) Small molecule targeting CELF1 RNA-binding activity to control HSC activation and liver fibrosis. *Nucleic Acids Res.* **50**, 2440–2451
67. Meisner, N. C., Hintersteiner, M., Mueller, K., Bauer, R., Seifert, J. M., Naegeli, H. U., *et al.* (2007) Identification and mechanistic characterization of low-molecular-weight inhibitors for HuR. *Nat. Chem. Biol.* **3**, 508–515
68. Wu, X., Ramesh, R., Wang, J., Zheng, Y., Armaly, A. M., Wei, L., *et al.* (2023) Small molecules targeting the RNA-binding protein HuR inhibit tumor growth in xenografts. *J. Med. Chem.* **66**, 2032–2053
69. Minuesa, G., Albanese, S. K., Xie, W., Kazansky, Y., Worroll, D., Chow, A., *et al.* (2019) Small-molecule targeting of MUSASHI RNA-binding activity in acute myeloid leukemia. *Nat. Commun.* **10**, 2691
70. Backman, T. W., Cao, Y., and Girke, T. (2011) ChemMine tools: an online service for analyzing and clustering small molecules. *Nucleic Acids Res.* **39**, W486–W491
71. Diarra Dit Konte, N., Krepl, M., Damberger, F. F., Ripin, N., Duss, O., Sponer, J., *et al.* (2017) Aromatic side-chain conformational switch on the surface of the RNA Recognition Motif enables RNA discrimination. *Nat. Commun.* **8**, 654
72. Clingman, C. C., Deveau, L. M., Hay, S. A., Genga, R. M., Shandilya, S. M., Massi, F., *et al.* (2014) Allosteric inhibition of a stem cell RNA-binding protein by an intermediary metabolite. *Elife* **3**, e02848
73. Radoux, C. J., Olsson, T. S., Pitt, W. R., Groom, C. R., and Blundell, T. L. (2016) Identifying interactions that determine fragment binding at protein hotspots. *J. Med. Chem.* **59**, 4314–4325
74. Loch, J. I., Barciszewski, J., Sliwiak, J., Bonarek, P., Wrobel, P., Pokrywka, K., *et al.* (2022) New ligand-binding sites identified in the crystal structures of beta-lactoglobulin complexes with desipramine. *IUCrJ* **9**, 386–398
75. Slabicki, M., Kozicka, Z., Petzold, G., Li, Y. D., Manojkumar, M., Bunker, R. D., *et al.* (2020) The CDK inhibitor CR8 acts as a molecular glue degrader that depletes cyclin K. *Nature* **585**, 293–297
76. Bradshaw, W. J., Kennedy, E. C., Moreira, T., Smith, L. A., Chalk, R., Katis, V. L., *et al.* (2024) Regulation of inositol 5-phosphatase activity by the C2 domain of SHIP1 and SHIP2. *Structure* **32**, 453–466.e6
77. Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., *et al.* (2015) GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1-2**, 19–25
78. Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., *et al.* (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem.* **24**, 1999–2012
79. Barraud, P., and Allain, F. H. (2013) Solution structure of the two RNA recognition motifs of hnRNP A1 using segmental isotope labeling: how the relative orientation between RRMs influences the nucleic acid binding topology. *J. Biomol. NMR* **55**, 119–138
80. Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., and Beckstein, O. (2011) MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **32**, 2319–2327
81. Krojer, T., Talon, R., Pearce, N., Collins, P., Douangamath, A., Brandao-Neto, J., *et al.* (2017) The XChemExplorer graphical workflow tool for routine or large-scale protein-ligand structure determination. *Acta Crystallogr. D Struct. Biol.* **73**, 267–278
82. Ng, J. T., Dekker, C., Kroemer, M., Osborne, M., and von Delft, F. (2014) Using textons to rank crystallization droplets by the likely presence of crystals. *Acta Crystallogr. D Biol. Crystallogr.* **70**, 2702–2718
83. Collins, P. M., Ng, J. T., Talon, R., Nekrosiute, K., Krojer, T., Douangamath, A., *et al.* (2017) Gentle, fast and effective crystal soaking by acoustic dispensing. *Acta Crystallogr. D Struct. Biol.* **73**, 246–255
84. Winter, G., Lobley, C. M., and Prince, S. M. (2013) Decision making in xia2. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1260–1273
85. Winter, G. (2010) xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190
86. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. Section D* **69**, 1204–1214
87. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., *et al.* (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877