OXFORD

Functional characterization of *HNF4A* gene variants identify promoter and cell line specific transactivation effects

Alba Kaci[1,2], Marie Holm Solheim[1], Trine Silgjerd[3], Jorunn Hjaltadottir[1,3], Lorentze Hope Hornnes[4], Janne Molnes[1,5], Andre Madsen[6],

Gry Sjøholt[3], Christine Bellanné-Chantelot[7], Richard Caswell [iD][8], Jørn V. Sagen[1,4], Pål R. Njølstad[1,9], Ingvild Aukrust[1,5],

Lise Bjørkhaug [iD][3,*]

[1]Mohn Center for Diabetes Precision Medicine, Department of Clinical Science, University of Bergen, Haukelandsbakken 1, Bergen 5020, Norway
[2]Center for Laboratory Medicine, Østfold Hospital Trust, Kalnesveien 300, Grålum 1714, Norway
[3]Department of Safety, Chemistry, and Biomedical Laboratory Sciences, Western Norway University of Applied Sciences, Inndalsveien 28, Bergen 5020, Norway
[4]Department of Medical Biochemistry and Pharmacology, Haukeland University Hospital, Jonas Lies veg 87, Bergen 5021, Norway
[5]Department of Medical Genetics, Haukeland University Hospital, Jonas Lies veg 87, Bergen 5021, Norway
[6]Department of Clinical Science, University of Bergen, Jonas Lies veg 87, Bergen 5020, Norway
[7]Département of Medical Genetics, Sorbonne University, AP-HP, Hôpital Pitié-Salpêtriére, 21 rue de l'école de médecine, 75006 Paris, France
[8]Exeter Genomics Laboratory, Royal Devon University Healthcare NHS Foundation Trust, Barrack Rd, Exeter EX2 5DW, United Kingdom
[9]Children and Youth Clinic, Haukeland University Hospital, Haukelandsbakken 1, Bergen 5021, Norway
*Corresponding author. Department of Safety, Chemistry, and Biomedical Laboratory Sciences, Western Norway University of Applied Sciences, Bergen 5020, Norway. E-mail: Lise.Bjorkhaug.Gundersen@hvl.no

Abstract

Hepatocyte nuclear factor-4 alpha (HNF-4A) regulates genes with roles in glucose metabolism and β-cell development. Although pathogenic *HNF4A* variants are commonly associated with maturity-onset diabetes of the young (MODY1; HNF4A-MODY), rare phenotypes also include hyperinsulinemic hypoglycemia, renal Fanconi syndrome and liver disease. While the association of rare functionally damaging *HNF1A* variants with HNF1A-MODY and type 2 diabetes is well established owing to robust functional assays, the impact of *HNF4A* variants on HNF-4A transactivation in tissues including the liver and kidney is less known, due to lack of similar assays. Our aim was to investigate the functional effects of seven *HNF4A* variants, located in the HNF-4A DNA binding domain and associated with different clinical phenotypes, by various functional assays and cell lines (transactivation, DNA binding, protein expression, nuclear localization) and *in silico* protein structure analyses. Variants R85W, S87N and R89W demonstrated reduced DNA binding to the consensus HNF-4A binding elements in the *HNF1A* promoter (35, 13 and 9%, respectively) and the *G6PC* promoter (R85W ~10%). While reduced transactivation on the *G6PC* promoter in HepG2 cells was shown for S87N (33%), R89W (65%) and R136W (35%), increased transactivation by R85W and R85Q was confirmed using several combinations of target promoters and cell lines. R89W showed reduced nuclear levels. *In silico* analyses supported variant induced structural impact. Our study indicates that cell line specific functional investigations are important to better understand HNF4A-MODY genotype–phenotype correlations, as our data supports ACMG/AMP interpretations of loss-of-function variants and propose assay-specific *HNF4A* control variants for future functional investigations.

Keywords: HNF4A-MODY; functional studies; transactivation; DNA binding; nuclear localization

Introduction

Hepatocyte nuclear factor-4 alpha (HNF-4A) is a transcription factor encoded by the *HNF4A* gene and plays a crucial role in the development, differentiation and function of the pancreas (β-cells), liver (hepatocytes), kidney (proximal tubules) and intestines (epithelium) [1–3]. It belongs to the nuclear receptor (NR) superfamily and comprises the following functional domains: transactivation domains AF-I and AF-II located in the N-terminus and C-terminus of the protein, a DNA-binding domain (DBD) including two zinc fingers and a hinge region, a dimerization/ligand-binding domain (LBD), and a C-terminal repressor domain [4]. The repressor domain, which inhibits the full transactivation potential of AF-II, is a unique feature among members of the NR superfamily. HNF-4A functions as a homodimer regulating the expression of target genes by a head-to-tail binding to specific and repetitive DNA response elements, referred to as DR1 [5, 6]. Heterodimerization by various HNF-4A specific isoforms, however, allows the regulation of a broader range of genes [7]. Further, the hydrophobic ligand binding pocket in the LBD domain facilitates the constitutive binding of fatty acids, allowing conformational change and interaction of co-activators [8] and subsequent HNF-4A activation.

Rare variants in the *HNF4A* gene have been associated with a variety of clinical phenotypes. Briefly, loss-of-function variants cause Maturity-Onset Diabetes of the Young type 1 (MODY1; HNF4A-MODY, OMIM #125850); a monogenic form of diabetes
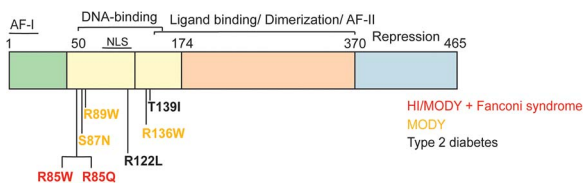
Figure 1. Schematic overview highlighting the HNF-4A protein sequence (isoform 2; NP_000448.3) and functional domains, as well as positions of HNF-4A variants investigated and associated with different glycemic phenotypes.

characterized by progressive pancreatic β-cell dysfunction and impaired insulin secretion [9]. Some variants, however, result in a biphasic phenotype manifesting with neonatal hyperinsulinemic hypoglycemia (HH) and HNF4A-MODY later in life [10]. Further, variants affecting the R85 residue in HNF-4A have been reported in isolated families presenting with an atypical form of Fanconi renotubular syndrome (OMIM #616026), characterized by inadequate reabsorption of glucose, amino acids and low-molecular weight proteins in the proximal renal tubules of the kidney [11]. Moreover, common variants in *HNF4A* have been associated with a moderate increased risk for the development of type 2 diabetes (T2D) [12, 13]. As there have been few studies investigating naturally occurring *HNF4A* variants for functional impact, an explanation for the diversity in *HNF4A* genotype-phenotypes remains largely unknown. The availability of suitable and robust control variants in functional assays for evaluating variants' effects for diagnostic purposes has therefore also been lacking. Due to this, there is a knowledge gap regarding *HNF4A* functional effect and its use in variant interpretation for precision medicine. This is in contrast to the more well characterized MODY3-associated *HNF1A* (HNF1A-MODY), for which guidelines have been developed to scale the functional impairment of variants [14]. To improve our understanding of the variant mechanistic effects underlying the *HNF4A* genotype-phenotype diversity, we hypothesized that this could be accomplished through investigations of the functional and structural consequences of seven variants located in the HNF-4A DBD and reported associated with either HNF4A-MODY, HNF4A-MODY and atypical Fanconi syndrome, or identified in rare T2D cases (Fig. 1 and Supplementary Table S1). Since HNF-4A isoform 2 (NM_000457.4), encoding a 474-amino acid protein, is the most highly expressed isoform in kidney and liver cells, variants in this isoform were investigated in terms of their transactivation activity using various promoter-linked reporters and cell line systems, for alterations in DNA-binding ability, and for protein expression and nuclear localization levels.

Aiming towards developing robust and meaningful functional HNF-4A assays, our study demonstrates the relevance of investigating promoter-specific transactivation in multiple relevant cell lines, supplemented with DNA binding and nuclear localization assays, to best evaluate *HNF4A* variant functional effects and to identify suitable assay control variants. Applying *in silico* structural tools for predicting possible variant induced structural effects was also a valuable supplement to our *in vitro* functional findings.

Results
Selection of variants

Five previously reported rare *HNF4A* missense variants classified as either variants of unknown significance (VUS) and including S87N and R122L, likely pathogenic (LP) and including

R85Q, or pathogenic (P) and including R85W and R89W (Supplementary Table S1) [15], were functionally investigated. The remaining two variants were selected on the basis of expectations of a milder functional effect as being reported as either a low-penetrant HNF4A-MODY variant (R136W) [16, 17] or classified as benign with respect to MODY (T139I) [18]. Moreover, the selection of variants was also based on being located in close proximity (R85W, R85Q, S87N, R89W) in the HNF-4A DBD sequence, however being associated with different glycemic phenotypes in variant carriers.

Altered transactivation of HNF-4A variants on *HNF1A* and *G6PC* promoters

The seven HNF-4A variants were first investigated for their ability to activate transcription in transiently transfected cell lines. HeLa cells were selected to avoid interference of endogenous HNF-4A. HNF-4A variant activity was investigated using an *HNF1A* promoter-linked luciferase reporter construct containing an HNF-4A recognition site (Fig. 2a). Although unable to reach statistical significance, R89W presented with lowest transactivation activity (< 50%), while S87N, R136W and R122L activities were ∼70%–80% compared to WT activity. Near-normal activity was observed for T139I. Interestingly, R85W trended towards increased activity (∼150%).

The relative protein expression levels of HNF-4A variants in cell lysates obtained from the transactivation assay (Fig. 2a) was further analyzed by SDS-PAGE and immunoblotting (Supplementary Fig. S1a–c). Normalization of activity obtained from the transactivation assay with variant protein expression levels did not dramatically alter the initial trend of the activity levels. Increased activity was still indicated for R85W (∼160%), while still reduced, however non-significant, for R89W and S87N (43% and 45%, respectively) (Fig. 2b). Transactivation and protein expression data are also summarized in Table 1.

To confirm variant activity on a different HNF-4A target promoter and relevant cell line, we assessed the *G6PC* promoter-linked luciferase reporter in transiently transfected HepG2 cells (Fig. 2c). The low background activity by the empty vector (EV) indicated minimal interference by endogenous proteins on this reporter. The R85W variant demonstrated increased activity also on the *G6PC* promoter (> 350%). Variants S87N and R136W, however, demonstrated strongly reduced activity (47%) on the *G6PC* promoter. Activities of R122L, R89W and T139I were either 60% (R122L) or > 75% (Fig. 2d). Of note, normalization of transactivation by HNF-4A protein expression levels in transfected HepG2 cells was not suitable due to strong presence of endogenous HNF-4A protein detected by immunoblot analysis when using an HNF-4A specific antibody (Supplementary Fig. S1d–f).

R89W and R85W differentially affects HNF-4A nuclear localization level

Subsequently, we investigated whether the five HNF4A-MODY-associated *HNF4A* variants could disturb normal nuclear localization of HNF-4A (Fig. 3, Table 1). For this, we used transiently transfected HeLa cells. The relative nuclear level was measured after cell fractionation (nuclear/cytosol), SDS-PAGE and immunoblotting using HNF-4A- and nuclear/cytosol marker specific antibodies (Supplementary Fig. S2). Quantification of HNF-4A in nuclear fractions demonstrated strongly reduced levels of R89W (∼20%), compared to WT (set to 100%), and increased levels of R85W and R85Q (∼240%- and ∼160%, respectively). All variants demonstrated exclusively nuclear localization (absent in cytosol) apart from the R89W variant, which was also partially detected in
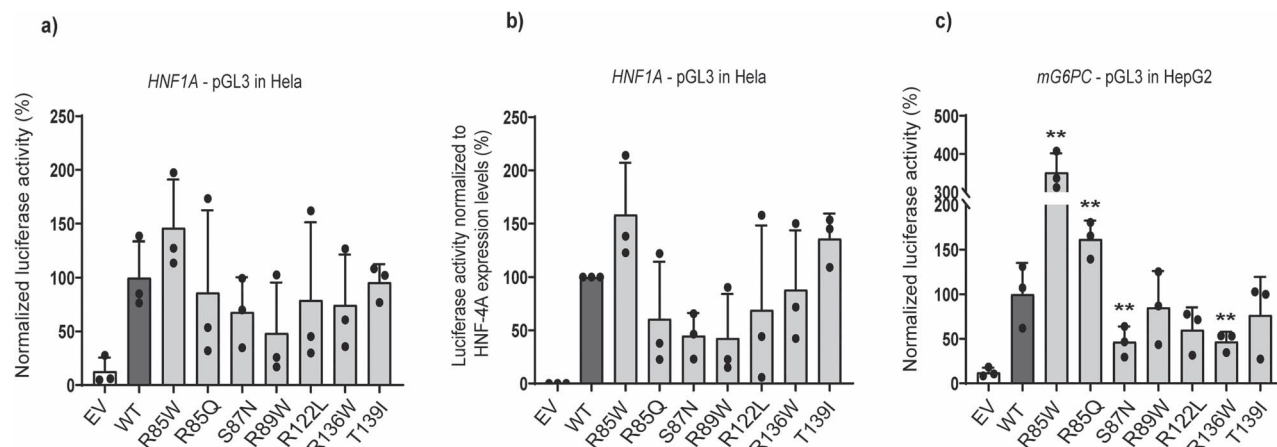
Figure 2. Transcriptional activity of HNF-4A variants investigated in transiently transfected (a) HeLa cells on pGL3-*HNF1A*, and in (c) HepG2 cells on pGL3-*G6PC* reporter plasmids. (b) Transcriptional activity after normalizing for protein expression levels in HeLa cells. For protein expression levels in HeLa/HepG2 see Supplementary Fig. S1. Empty vector (EV) was used as a negative control. Each bar represents the mean of nine readings ± SD; taken from three parallel readings from each of each of three independent experimental days. *indicates $P < 0.05$.

Table 1. Summary of functional investigation data on HNF-4A variants. The PS3_supporting or BS3_supporting ACMG scores were only applied when a variant was shown to be impaired (transactivation activity < 60% and DNA binding < 60%, compared to wild type) or showed no damaging effect (transactivation activity > 75% of wild type), respectively, in at least two functional assays by transactivation and/or DNA binding assay. The nomenclature used for the variants is based on the *HNF4A* transcript NM_000457.4 (HNF4a isoform 2). Values given in % of WT (100%); N/A, not available; N/R, not relevant.

HNF-4A variant	Transactivation *HNF1A prom* (HeLa)	Protein expression (HeLa)	DNA binding *HNF1A* (HeLa)	Nuclear localization (HeLa)	Transactivation *G6PC prom* (HepG2)	DNA binding *G6PC* (HepG2)	ACMG supporting functional evidence
EV	20	0	7	0	12	5	N/R
WT	100	100	100	100	100	100	N/R
R85W	146	92	35	240	352	10	-
R85Q	86	142	75	160	162	54	-
S87N	68	151	13	80	47	N/A	PS3_supporting
R89W	48	114	9	20	85	N/A	PS3_supporting
R122L	79	103	N/A	N/A	60	N/A	-
R136W	74	85	94	94	47	N/A	-
T139I	96	70	N/A	N/A	77	N/A	BS3_supporting

the cytosolic fraction by multiple lower molecular weight bands (Supplementary Fig. S2).

HNF4A variants severely affects HNF-4A DNA binding

As the five HNF4A-MODY-associated *HNF4A* variants are located in the HNF-4A DBD, we investigated their effect on HNF-4A binding to an HNF-4A recognition site present in the *HNF1A*- and *G6PC* short oligonucleotide sequences (Fig. 4a, Table 1). All variants except R85Q and R136W demonstrated reduced DNA binding to the recognition site present in the *HNF1A* promoter; DNA binding was severely reduced for variants S87N, R89W and R85W (~13%-, 9% and 35%, respectively) (Fig. 4b, Supplementary Fig. S3a). A supershift assay further verified the presence of HNF-4A variant proteins in the *HNF1A*-oligo bound complexes (Supplementary Fig. S3b). However, none of the variants tested appeared to exert any dominant negative effect on DNA binding by WT HNF-4A (Supplementary Fig. S3c). For a follow-up investigation of the differential effect of the two variants (R85W and R85Q) affecting the same residue (R85), we assessed their binding to the *G6PC* oligonucleotide sequence, for interrogating the consequence of R85 making fewer base pair contacts within the HNF-4A

recognition sequence of *G6PC* [6] compared to the *HNF1A* oligo (Fig. 4b). Variant R85W showed significantly reduced binding to the *G6PC* oligo (~10%), similar to the effect on binding to the *HNF1A* oligo, while R85Q, which showed near normal binding to the *HNF1A* oligo, demonstrated only ~50% binding to the *G6PC* oligo.

R85W and R85Q differentially affect promoter transactivation in different cell lines

Intrigued by our findings concerning R85W in particular, which poorly binds the HNF-4A recognition site in the *HNF1A*- and *G6PC* short oligonucleotide sequences (Fig. 4b and c), but demonstrated increased transactivation of the reporters driven by the larger *HNF1A*- and *G6PC* promoter sequences in HeLa and HepG2 cells, respectively (Fig. 2a–c), we aimed to investigate whether R85W activity on these subsequent reporters was also increased in additional relevant cell lines. For this, transactivation was also assessed in HK2, HepG2 and MIN6 cells (Fig. 5 and Table 1). For comparison we included R85Q, which affect the same residue as R85W but had a less deleterious impact in DNA binding and HeLa/HepG2 transactivation assay (Fig. 4b and c, Fig. 2a–c). In transiently transfected HK2 cells (Fig. 5a), R85W significantly
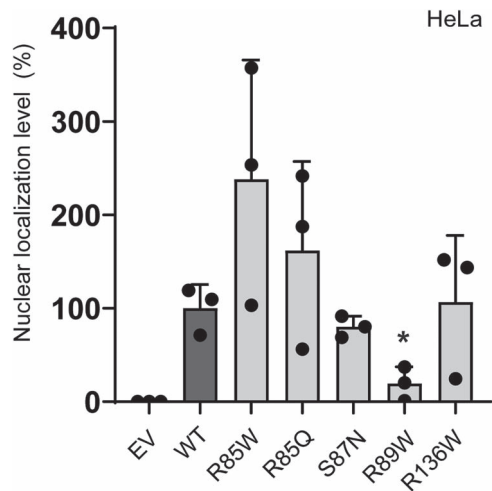
Figure 3. Effect of selected *HNF4A* variants on the nuclear localization of HNF-4A. Nuclear and cytosolic fractions were analyzed by SDS-PAGE and immunoblotting. Each bar represents the mean level of HNF-4A variants in nuclear fractions from three-four independent nuclear fractionation experiments (n = 3–4). *indicates *P* < 0.05. EV = empty vector. WT = wild type. Representative western blots are presented in Supplementary Fig. S2.

increased activity on the *HNF1A* promoter compared to WT (~650% activity) and this increase was greater than that observed in HeLa cells. R85Q also showed increased activity by ~260% compared to WT in HK2 cells, whereas there was no significant difference in HeLa cells. Surprisingly, in HepG2 and MIN6 cells, both R85W and R85Q presented with slightly reduced/near normal activity (~75%–80%) on the *HNF1A* promoter (Fig. 5b and c). The combination of *HNF1A*/HepG2 cells, however, indicates challenges concerning interference of presumably endogenous HNF-4A on the *HNF1A* promoter, shown by higher activity in the EV sample (> 35%) (Fig. 5c), compared to that detected for the HNF1A/HK2 and HNF1A/MIN6 combinations (Fig. 5a and b).

We also investigated variant activity on the *G6PC* promoter-linked reporter in HK2 cells (Fig. 5d). Results indicated similar activity as in HepG2 cells, by increased activity of R85W (~180%) and near normal activity of R85Q. Overall, our transactivation assays indicate a cell line specific effect of variant activity on particularly the *HNF1A* promoter (Fig. 2, Fig. 5). As *G6PC* is not expressed in pancreatic *β*-cells, transactivation of HNF-4A variants was not assessed in MIN6 cells.

We further wanted to see whether we could confirm the variant specific effect on endogenous *G6PC* and *HNF1A* regulation and expression. For this, RNA was extracted from HK2 cells transfected with expression constructs for HNF-4A variants, and *G6PC* and *HNF1A* expression was investigated by real-time qPCR (Supplementary Fig. S4). A modest increase in *G6PC* expression was detected for R85W (1.7-fold versus WT of 1.6-fold), while a significantly reduced expression of *G6PC* was seen by R85Q (1.4-fold). In contrast, the variants had no significant effect on the expression of endogenous *HNF1A*.

In silico analyses of HNF-4A variants indicate variant induced structural effects

In silico structural analyses were performed on HNF-4A variants aiming for a structural explanation for variant induced loss-of-function or gain-of-function detected by our *in vitro* functional assays. Various software tools were applied for predicting the residue change effect on the protein flexibility, DNA

interaction and binding through available hydrogen bonds, the entropy and energy change, and the stability of the HNF-4A variant protein (Supplementary Tables S2–S5). Only variant residues located within the crystallized protein structure (PDB i.d. 3CBB) were analyzed (Fig. 6a; five of seven variants). Changes in molecular flexibility by residue induced protein structure deformations, as analyzed by Eris, are shown in Fig. 6b. Illustration of residue induced protein structure deformations, as predicted by DynaMut, is shown in Fig. 6c.

Compared to R85, there was strong consensus among the software tools used that W85 was predicted to be stabilizing for the HNF-4A protein structure (Supplementary Table S2). Of all amino acid residues studied, W85 was also predicted to most impact the flexibility (reduced) of the molecular structure (helix residues 71–92 and residues Q131-E133), and the change in free vibrational energy ($\Delta\Delta S$ Vib). A reduction in available hydrogen bonds for DNA-binding by W85 was also predicted, as well as a reduction in free binding energy, compared to R85. Variant residue Q85 (compared to R85) was also predicted to reduce the flexibility of the HNF-4A helix structure from residues 85–89 and residues N132 and E133 (increased for V130 and Q131), as well as reducing the number of hydrogen bonds available for DNA-binding, however not to the extent as W85. Although variant residue W89 most strongly impacted the free binding energy of residue 89 (compared to wild type residue R89), W89 was also predicted to cause a deformation of the molecular structure including residues Q111 and C112 (Fig. 6c). Further, a complete loss of hydrogen bonds was also predicted for W89. The S87 to N87 residue change was not predicted to affect the number of direct hydrogen bonds, as none are in direct contact with DNA, and was rather predicted to result in an additional indirect hydrogen bond. Further, increased molecular flexibility of the protein was predicted for the S87N variant, particularly for residues 78–89 within the same helix as the substitution, and for residues M93 to Q131. Reduced flexibility was predicted for residues L90-H92 and E133. The predicted effects of R85W, R85Q, R89W and S87N on protein structure/stability were also supported by their high REVEL scores of 0.949, 0.921, 0.913 and 0.925, respectively (Supplementary Table S1), and further supporting a likely deleterious impact of the HNF-4A variant.

Although the molecular flexibility was increased for R122, no effect on the free energy of binding was predicted for the L122 to R122 change, supported also by the lower REVEL score of R122L (0.776) compared to the other HNF-4A protein variants investigated by structural *in silico* analyses (Supplementary Table S1). Further, although the L122R substitution was predicted to result in loss of two hydrogen bonds, this amino acid sidechain faces away from DNA in the recognition helix and so the variant has no direct impact on DNA-binding.

Discussion

Investigation of functional impact of MODY-associated gene variants by *in vitro* functional studies is an important supplement in evaluating the degree of variant-induced protein dysfunction and disease causality, and to disclose clinical actionable variants [19, 20]. The well characterized *HNF1A* gene and HNF-1A protein has largely contributed to this, where variant molecular pathogenicity is defined by sets of threshold values in functional assays and specified by the gene specific ACMG/AMP variant interpretation guidelines by the ClinGen Monogenic Diabetes Expert Panel [21]. According to the Human Gene Mutation Database, 202 disease causing variants (> 115 missense) have so far been reported in
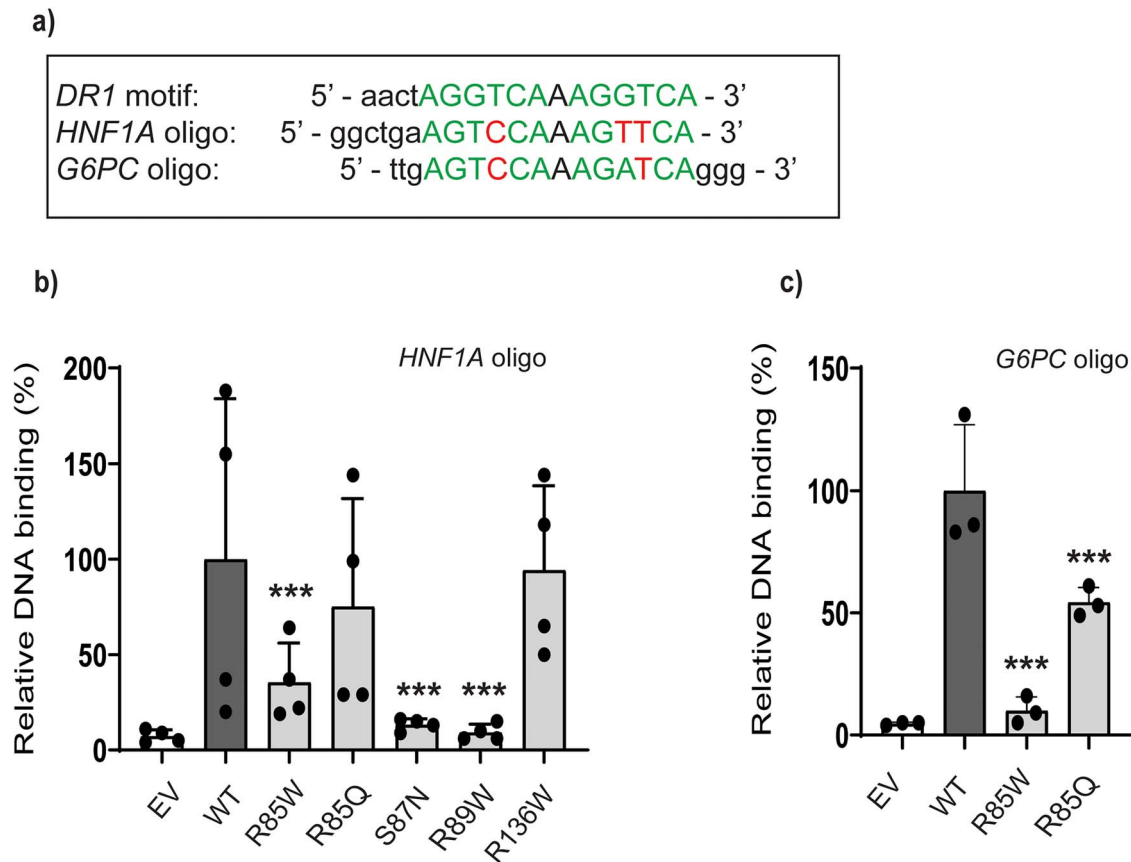
a)

DR1 motif:	5' - aactAGGTCAAAGGTCA - 3'
HNF1A oligo:	5' - ggctgaAGTCCAAAGTTCA - 3'
G6PC oligo:	5' - ttgAGTCCAAAGATCAggg - 3'

b)

c)

Figure 4. DNA binding of selected *HNF4A* variants by electrophoretic mobility shift analysis (EMSA). (a) HNF-4A recognition site within the *HNF1A* and *G6PC* promoter-based oligo sequences are shown, and relative to the DR1 element consensus recognition site. HNF-4A site base-specific nucleotide contacts made by R85 in *HNF1A* according to [6] and relative to the *G6PC* oligo, are also outlined. HNF-4A variants binding to the (b) *HNF1A* and (c) *G6PC* oligos. Following EMSA, bound complexes were quantified by densitometric analysis. Each bar represents the mean of three independent EMSA experiments (n = 3) ± SD. **indicates *P* < 0.01. EV = empty vector. WT = wild type. Representative gel images are presented in Supplementary Fig. S3.

the *HNF4A* gene. As only a few variants located in the protein coding region have been functionally studied to date [22], the basis for guiding threshold values for functional evaluation of variant induced HNF-4A dysfunction, for diagnostic purposes, is currently lacking. To tackle this, increased knowledge of the gradient of *HNF4A* variant functional effect is needed. Such studies should be based on robust and sensitive *in vitro* functional assays, in relevant cell line systems, that can distinguish between variants associated with different clinical phenotypes.

The maintenance of glucose homeostasis involves several complementary physiologic processes including glucose resorption (GI tract) and reabsorption (kidneys), glycogenolysis (liver), gluconeogenesis (liver/kidney) and glucose excretion (kidney). In an attempt to capture an *HNF4A* variant effect in multiple systems, our functional HNF-4A investigations were performed in several cell lines, three of which represent cell models from tissues contributing to glucose homeostasis. Further, the *HNF1A* and *G6PC* genes are known HNF-4A targets for DNA binding and transactivation in tissues where HNF-4A plays a central gene regulatory role, like for instance the kidney, liver and pancreas [23].

Our study demonstrates the relevance of the *in vitro* DNA binding assay for determining effects of variants located in the HNF-4A DBD on binding to recognition sites in *HNF1A* and *G6PC* promoters (Fig. 4). In binding assays using the *HNF1A* short oligonucleotide, we observed clear differences between the impact of variants S87N and R89W compared to that of R136W. A severe impact of R85W and R89W variants on binding to the ApoA1 site A probe has

previously been reported [24], while we observed varying degrees of reduced binding by variants at R85 to the *G6PC* oligonucleotide. The low DNA binding exhibited by variant S87N is in contrast to a report by Chandra et al. [25], which in a crystal structure of HNF-4A DBD-hinge-LBD multi-domain bound to DNA, identified S87 as facing away from DNA within the recognition helix and thus unable to participate directly in DNA binding, as confirmed by our *in silico* structural analyses. However, S87 has been shown to undergo targeted phosphorylation by protein kinase C, and this modification disrupted the ability of HNF-4A to bind DNA [26]. Residues R85 and R89, on the other hand, have been reported to be in direct contact with the DNA response element AGGTCA half-sites [6, 25]. This is in line with our findings of severely reduced DNA binding by variants R85W and R89W. Our *in silico* analyses also predicted W85 and W89 to dramatically reduce the number of hydrogen bonds (also reduced for Q85, Supplementary Table S4). The reduced binding of variants R85W and R85Q to the *G6PC* oligo compared to the *HNF1A* oligonucleotide could be a consequence of R85 making three base pair contacts within the HNF-4A recognition site of the *HNF1A* oligo, versus two in the *G6PC* oligo [6]. The near-normal DNA binding of R136W to *HNF1A* is, however, somewhat in contrast to previous reports investigating binding to *DR1/HNF1A* sequences, using purified HNF-4A protein or from nuclear extracts [25, 27]. Although R136W was not assessed for *G6PC* binding in the DNA binding assay, a loss-of-transactivation function (< 50%) was, however, found on the *G6PC*-reporter in HepG2 cells (Fig. 2). This could suggest a potential stronger variant
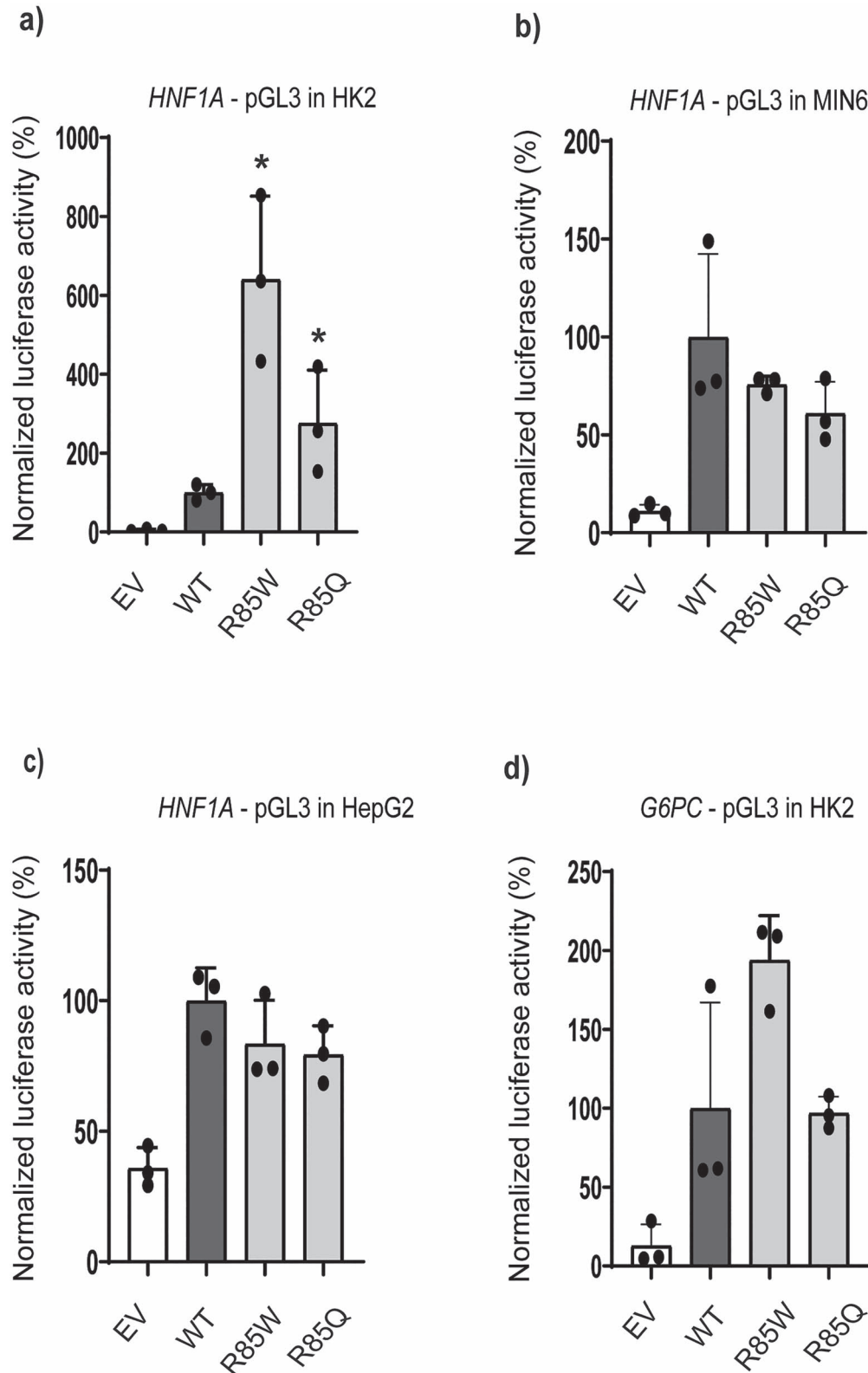
Figure 5. Transactivation by HNF-4A variant on *HNF1A*- and *G6PC*-linked reporter systems in transiently transfected (a and d) HK2, (b) MIN6, and (c) HepG2 cells transfection by WT or *HNF4A* variant plasmids together with (a–c) pGL3-*HNF1A* and (d) pGL3-*G6PC* reporter plasmids. Empty vector (EV) was used as a negative control. Each bar represents the mean of nine readings ± SD, taken from three parallel readings from each of three independent experimental days. *indicates $P < 0.05$.
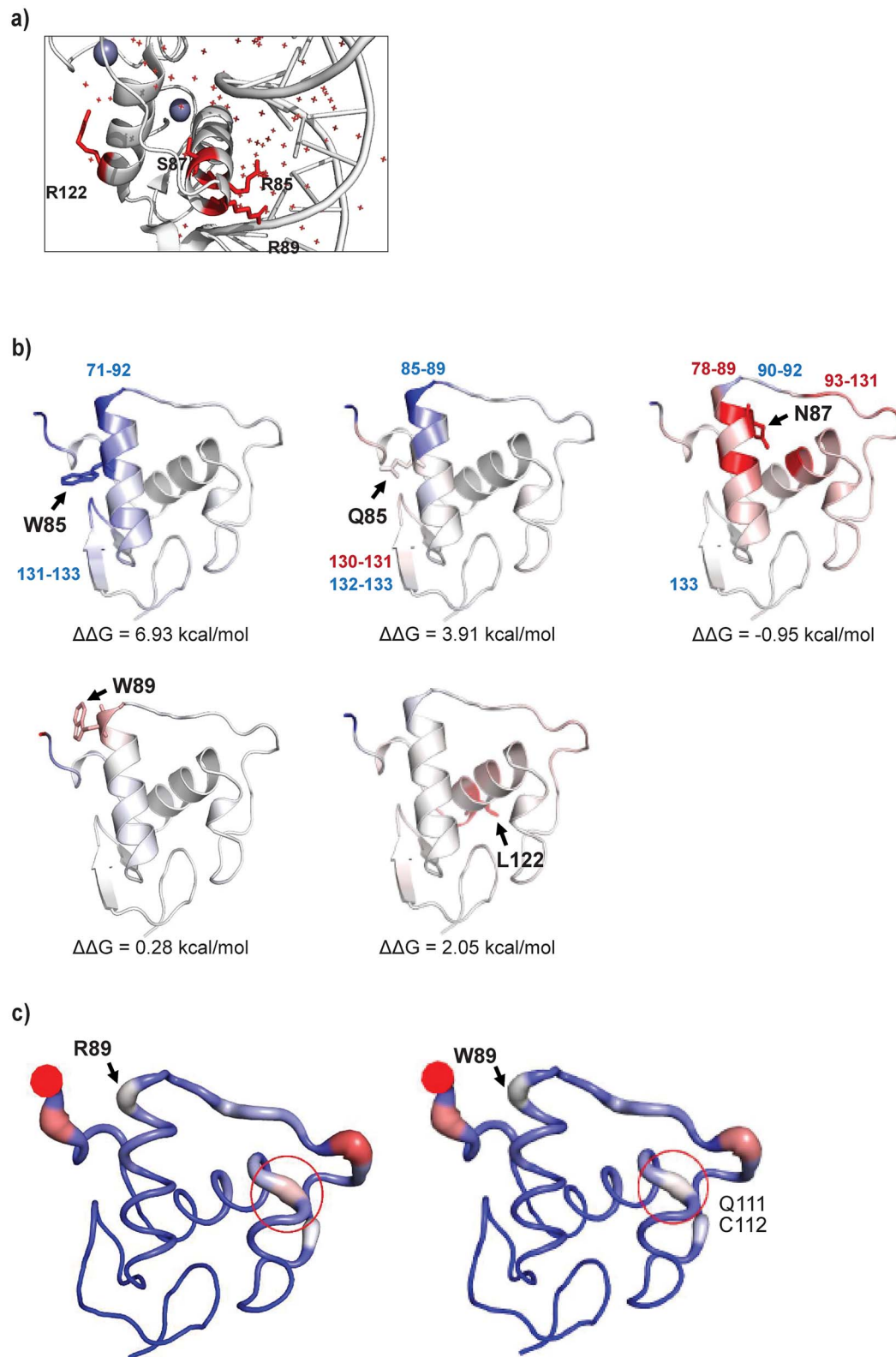
Figure 6. Protein structural changes induced by HNF-4A variant residues. (a) Position of HNF-4A residues subject to change. (b) Protein flexibility change induced by variants according to Eris. Changes in vibrational free energy are indicated by color: Blue indicates increased molecular rigidity, while red indicates increased flexibility compared to WT. R85W: Reduced flexibility of helix residues 71–92 and for Q131-E133. R85Q: Reduced flexibility for helix residues 85–89. S87N: Increased flexibility of helix residues 78–89 and residues M93-Q131, reduced flexibility of L90-H92 and E133. Position of variant residues are shown by arrows. (c) Deformation of protein structure by R89W as predicted by DynaMut. Protein deformation or strain indicated by widening of the molecular backbone and (highlighted circle), reduced strain/deformation predicted at residues Q111 and C112.

effect on *G6PC*- compared to *HNF1A* transactivation. R136 has further been reported to lie in a sensitive domain-domain junction, forming domain-domain arrangements with the upstream DBD in the HNF-4A crystal structure [25].

HeLa cells have been shown to be a suitable cell line for studying the functional consequence of *HNF1A* variants on HNF-1A transactivation and for successfully distinguishing pathogenic variants from benign and T2D associated variants [28]. In our study, a functional discrimination between severe (HNF4A-MODY linked) and milder (T2D associated) *HNF4A* variants on HNF-4A transactivation (*HNF1A* reporter) in HeLa cells was less evident and clouded by substantial intra-variant variations. Use of the *G6PC*- linked reporter system in HepG2 cells, however, was more successful in distinguishing between near-normal, loss- and gain-of-transactivation function variants (Fig. 2). The near-normal activity of the T2D associated variant T139I is in line with that reported elsewhere on *HNF1A*- or *TK*-linked reporter constructs [29, 30]. The high transactivation activity of R85W on both *G6PC*- and *HNF1A* promoter-linked reporters in HepG2 and HK2 cells (partly also for R85Q variant), is intriguing (Figs 2 and 5), particularly considering the low DNA binding ability (Fig. 4). It is important thus to bear in mind the differences in the size of the promoter-specific DNA sequences used in the DNA binding versus transactivation assays. The interaction of HNF-4A with other promoter-bound factors could potentially allow recruitment of HNF-4A to the promoter-linked transactivation reporters even in the presence of reduced direct binding to DNA. A contribution of potential HepG2 and HK2 cell-specific coactivators to the increased R85W promoter activity would, however, require more detailed investigations in a future study. Worth noting, however, is that in HepG2 cells, the peroxisome proliferator-activated receptor (PPAR)γ coactivator 1-alpha (PGC-1α) has been shown to coactivate HNF-4A, and that the presence of adjacent RORα/SRC-2 is necessary for the full transcriptional effect of PGC-1α [31]. Whether R85W, predicted to stabilize HNF-4A and reduce the flexibility of the molecular structure, forms a more stable interaction with PGC-1α in the HNF-4A C-terminal AF-II domain, is unknown. In addition to HNF-4A, these cofactors are also reported to be expressed at high levels in kidney tissue (the Human Protein Atlas).

Further, W85 was predicted to increase the rigidity and decrease the flexibility of helix 1 by our *in silico* analyses. In the absence of crystal structures of the HNF-4A AF-I and repressor domains, it is however impossible to predict how a putative increase in rigidity due to the R85W substitution might alter the potency of the AF-I domain activity function, or alternatively whether the increased activity of the variant results from an effect on the inhibitory function of the HNF-4A repressor domain [4, 32].

HNF-4A has been shown to play an important role in the upregulation of membrane transporters and receptor proteins for the reabsorption of various molecules in kidney proximal tubular epithelial cells [33], whereby the transport of glucose in and out of cells is mediated by specialized transport proteins [34]. Glucosuria, which is more frequent in HNF1A-MODY patients, is a biochemical feature also reported in R85W carriers [10]. Glucosuria may be due to a disturbance in glucose resorption or release by proximal tubular cells. The proximal tubule generates glucose-6-phosphate from various precursors, and through glucose-6-phosphatase subsequently generates free glucose that can exit the cell. The increased activity of R85W on the *G6PC*-linked reporter in HK2 cells (Fig. 5) may suggest increased regulation of glucose-6-phosphatase and increased glucose release by

kidney tubules (renal gluconeogenesis is projected to potentially be responsible for approximately 40% of all gluconeogenesis [35]). Further, increased activity of R85W on the same promoter in HepG2 cells (Fig. 2) may also suggest a mechanism of increased output of glucose from the liver.

There is much evidence for a cross-regulatory loop taking place between HNF-1A and HNF-4A in pancreatic cells [35, 36]. Stimulated circuits, very much like the HNF-1A/HNF-4A loop, have been previously shown to exhibit bistability [37, 38], where loss of one functional allele can increase the probability that the opposite gene is inhibited sufficiently to trigger the transition to an "OFF" state. Whether the mildly reduced R85W and R85Q activity on the *HNF1A*-linked reporter in MIN6 cells (Fig. 5) is sufficient to affect the bistability of such an HNF-1A/HNF-4A loop circuit in β-cells, causing reduced expression of HNF-1A and HNF-4A regulated downstream genes, requires more detailed investigations. Further, whether the biphasic phenotype of HH followed by HNF4A-MODY in association with the R85W variant could also be due to a contribution of increased gluconeogenesis in the kidneys and glycogenolysis from the liver in the post-absorptive state, is an interesting hypothesis that would also require further detailed investigations.

Our nuclear localization assay successfully separated variants with severely reduced nuclear levels (R89W, present at 20% of the normal level) from those with near-normal or increased levels (R85W and R85Q) (Fig. 3). The affected residue R89 has been reported to lie within the nuclear localization signal (NLS) of HNF-4A (residues 80–121), which also contains a site for acetylation and is important for proper nuclear retention [39]. Although our *in silico* analyses were inconclusive regarding a destabilizing effect of the R89W substitution (Supplementary Table S2), W89 was predicted to cause a deformation of the molecular structure of nearby residues, which might contribute to reduced levels of protein (Fig. 6). However, while NLSs are typically rich in lysine and arginine residues, the efficient nuclear localization observed for the R85W and R89W variants suggest that the effect of R89W is not simply due to a reduction in positive charge. Finally, the normal nuclear localization of R136W correlates with previous reports [27].

Since the criteria for evaluating variant molecular pathogenicity by functional assays, set by the ClinGen Monogenic Diabetes Expert panel and *HNF4A* specific AMCG/AMP guidelines, are directed towards loss-of-function variants, these guidelines are not applicable for interpretation of variants displaying assay specific gain-of-function (e.g. R85W and partly R85Q). Our functional data however supports T139I representing a benign variant since it qualifies for the use of BS3_supporting criteria according to the thresholds set by the *HNF4A* specific ACMG guidelines and R89W representing a pathogenic variant qualifying for the use of PS3_supporting criteria. Moreover, our functional data also supports pathogenicity for the S87N variant (PS3_supporting), however this variant still remains a VUS according to the ACMG guidelines (Table 1, Supplementary Table S1).

In conclusion, we recommend the use of the *G6PC*-linked reporter system in HepG2 cells to best capture the gradient of functional effect of *HNF4A* diabetes-associated variants. Moreover, as MODY gene variant investigations require appropriate internal controls for evaluating variant pathogenic effects by functional assays, our study identified the following control variants for studies in HepG2 cell transactivation assay; S87N for severely reduced activity (< 50%) and T139I for near-normal activity. Further, variant R89W is proposed as control variant for severely reduced nuclear localization in HeLa cells (< 25%),

and S87N/R89W as control variants for DNA binding using the *HNF1A* oligo (< 20%). The systematic use of the same control variants across laboratories in functional investigations of *HNF1A* variants has proven extremely important for capturing the range of variant effects across the *HNF1A* allelic spectrum. This has allowed recommendations of assay-specific threshold values for diagnostic purposes, defining what is "decreased" and "normal" function, and should also be implemented for future *HNF4A* variant investigations.

Materials and methods
HNF4A gene variants

Seven previously reported *HNF4A* missense variants (R85W, R85Q, S87N, R89W, R122L, R136W and T139I) associated with different glycemic phenotypes in variant carriers were selected for functional evaluation [15–18]. Their domain location, ClinVar and REVEL scores, and associated clinical phenotype, are shown in Supplementary Table S1. Further, *HNF4A* variant interpretation was performed according to the ClinGen Monogenic Diabetes Expert Panel specifications to the American College of Medical Genetics (ACMG) and Genomics/the Association for Molecular Pathology (AMP) guidelines for *HNF4A* variants, except for variant R136W which due to its reduced penetrance could not be classified by the standard ACMG guidelines. The PS3_supporting or BS3_supporting ACMG scores were only applied when a variant was shown to be functionally impaired (transactivation activity < 60% and DNA binding < 60%, compared to wild type) or showed no damaging effect (transactivation activity > 75% compared to wild type), respectively, in at least two functional assays by transactivation and/or DNA binding assay. The nomenclature used for the variants is based on the *HNF4A* transcript NM_000457.4 (HNF4a isoform2).

Constructs

All *HNF4A* variants were constructed using the QuikChange II XL Site-Directed Mutagenesis Kit (Agilent Technologies) and variant-specific primers (sequences available upon request). Individual *HNF4A* variants were introduced into the WT *HNF4A* cDNA isoform 2 (NCBI NM_000457.4), in FLAG-tagged pcDNA3.1+ vector. In the transactivation assays, the following reporter plasmids coupled to *Firefly luciferase* by *luc +* in pGL3 (Promega) were used: the human *HNF1A* gene promoter sequence (nucleotides −129 to +196) kindly provided by K. Yamagata (Osaka University, Japan) [27] or the mouse *G6PC* gene promoter sequence (nucleotides −231 to +66) kindly provided by B.W. O'Malley (Baylor College of Medicine, USA), and the *Renilla luciferase* control vector pRL-SV40 (Promega).

Cell lines and transfection

HeLa cells (human cervical carcinoma cells) were cultured in DMEM medium (without pyruvate) with 10% FBS. HepG2 cells (human hepatoma cells) were cultured in EMEM medium (with pyruvate) with 10% FBS and L-glutamine (290 mM). HK2-cells (human proximal tubule cells) were cultured in Keratinocyte Serum Free Medium (K-SFM) with 0.05 mg/ml Bovine pituitary extract and 5 ng/ml human recombinant Epidermal growth factor. MIN6 cells (mouse insulinoma cells) were cultured in DMEM medium with 15% FBS. All media were supplemented with penicillin/streptomycin (100 U/ml) and cell lines cultured at 37°C with 5% CO_2 in a humid atmosphere. For transfection purposes, HeLa, MIN6 and HK2 cells were transfected using Lipofectamine2000 (Invitrogen), while HepG2 cells were transfected using the TransIT-LT1 transfection reagent (Mirus Bio).

Luciferase assays and protein abundance

Cells were transiently transfected with wild-type (WT) or variant *HNF4A* plasmids, together with a *Firefly* reporter plasmid and the *Renilla* reporter plasmid. Luciferase activities were measured 24 h post transfection using the Dual-Luciferase Assay System (Promega) on a Centro XS3 LB 960 luminometer (Berthold Technologies). HNF-4A activity was normalized to *Firefly* and *Renilla* activity. The level of HNF-4A protein expression (corresponds to the canonical protein sequence P41235-1 in the UniProtKB entry for HNF-4A) of WT and HNF-4A variants was assessed in various cell lysates obtained for the transactivation assays. In short, 20 μl of cell lysates was subjected to SDS-PAGE and immunoblotting using specific antibodies against HNF-4A (Cell Signaling; #C11F12) and β-actin (Santa Cruz; sc-47778).

Subcellular fractionation

Nuclear and cytosolic fractions were isolated from transiently transfected HeLa cells as previously described [40]. Total protein quantification in each fraction was measured using the Pierce™ BCA Protein Assay Kit (Thermo Fisher) and 8–10 μg total protein from fractions was subjected to SDS-PAGE and immunoblotting using rabbit anti-HNF-4A (Cell Signaling; #C11F12). The purity of the fractions was verified using antibodies against the nuclear marker protein Topoisomerase II-α (Cell signaling; #12286) and the cytosolic marker protein GAPDH (Santa Cruz; sc-47724). The relative subcellular localization based on each fraction was determined by calculating the ratios of HNF-4A with the respective nuclear or cytosolic markers.

DNA binding assay

The electrophoretic mobility shift assay (EMSA) was carried out as previously described [36]. Briefly, equal amounts of proteins from nuclear fractions of transiently transfected HeLa cells were incubated together with cyanine 5-labelled oligonucleotides (Sigma Aldrich) and the binding reaction was performed using the Odyssey EMSA buffer kit (LI-COR Biosciences). The double stranded DNA fragments containing the HNF-4A binding site in the promoter of the *G6PC* gene (5′-TTGAGTCCAAAGATCAGGG-3′) or the HNF-4A binding site in the promoter of the *HNF1A* gene (5′-GGCTGAAGTCCAAAGTTCA-3′) were used in the binding reactions. Bound complexes were analyzed by electrophoresis on 6% (w/v) polyacrylamide gels for EMSA (Thermo Fisher).

RNA isolation, reverse transcription and real-time quantitative PCR (RT-qPCR)

Total RNA from HK2 cells, transiently transfected with WT or *HNF4A* variant plasmids, was extracted using the PureLink RNA Mini Kit (Thermo Fisher). RNA (300 ng) was reverse transcribed using the qScript™ cDNA Synthesis Kit (Quantabio, Beverly, MA). The effect of transfected *HNF4A* variants on target gene expression was analyzed by RT-qPCR using gene specific primers (sequences available upon request) and SYBRGreen Fast Mix kit (Thermo Fisher). Analysis was performed using theStepOnePlus™ Real-Time PCR system instrument (Thermo Fisher). The reference gene used was *RPL13*.

Structural analyses

Structural protein modelling of HNF-4A variants was carried out using the crystal structure of the DNA binding domain (residues

33–113; PDBC: 3CBB) and the side-chain rotamer library incorporated within the mutagenesis wizard of the PyMOL version 2.2.2 software (Schrödinger, Inc.). PyMOL, thus, also predicts hydrogen bond status of residues. DynaMut was used to predict the effect of variants on protein flexibility, interaction and stability, by the change in entropy ($\Delta\Delta G$) and vibrational energy ($\Delta\Delta S$) of the molecule (Supplementary Table S3) [41]. Eris, which is a protein stability prediction tool, calculates the change $\Delta\Delta G$ using the Medusa Modelling suite, and also models backbone flexibility [42]. Dezyme evaluates changes in the free energy of folding of the variant protein, as well as solvent accessibility [43]. MUpro [44] and CUPSAT [45] are predictors of changes in protein stability caused by missense variants. PreHot uses binding energy calculations to predict hotspots of regions crucial for the interaction between protein and DNA (Supplementary Table S5) [46]. SAMPDI predicts protein variant effect on DNA binding properties, also based on free energy of binding [47].

REVEL analysis

The Alamut Software (version 2.15, SOPHiA GENETICS) was used to extract REVEL scores for the HNF-4A variants. REVEL [48] is an ensemble method for predicting the damaging effect of missense variants to protein structure or function based on combining scores from 13 individual *in silico* tools: MutPred, FATHMM v2.3, VEST 3.0, SIFT, PolyPhen-s, PROVEAN, MutationTaster, MutationAssessor, LRT, GERP++, phyloP, SiPhy, and phastCons. REVEL scores range from 0 to 1, with higher scores reflecting greater likelihood that a variant is damaging/pathogenic (> 0.7).

Statistical analysis

All data are presented as mean ± standard deviation (STD) relative to WT (set as 100% or 1-fold), unless otherwise specified. Each experiment was carried out on three independent experimental days unless otherwise specified. Differences between the studied variants were analyzed using GraphPad Prism Software (version 8.1.1) and raw data (i.e. *Firefly/Renilla* ratios). Statistical analyses were performed using a 1-way ANOVA and Dunnett's posthoc test (alpha level of 0.05).

Supplementary data

Supplementary data is available at *HMG Journal* online.

Conflict of interest statement: The authors have no conflicts of interest to declare.

References

1. Chen WS, Manova K, Weinstein DC. *et al.* Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos. *Genes Dev* 1994;**8**:2466–77.

2. Stoffel M, Duncan SA. The maturity-onset diabetes of the young (MODY1) transcription factor HNF4alpha regulates expression of genes required for glucose transport and metabolism. *Proc Natl Acad Sci USA* 1997;**94**:13209–14.

3. Duncan SA, Navas MA, Dufort D. *et al.* Regulation of a transcription factor network required for differentiation and metabolism. *Science* 1998;**281**:692–5.

4. Hadzopoulou-Cladaras M, Kistanova E, Evagelopoulou C. *et al.* Functional domains of the nuclear receptor hepatocyte nuclear factor 4. *J Biol Chem* 1997;**272**:539–50.

5. Fang B, Mane-Padros D, Bolotin E. *et al.* Identification of a binding motif specific to HNF4 by comparative analysis of multiple nuclear receptors. *Nucleic Acids Res* 2012;**40**:5343–56.

6. Lu P, Rha GB, Melikishvili M. *et al.* Structural basis of natural promoter recognition by a unique nuclear receptor, HNF4alpha. Diabetes gene product. *J Biol Chem* 2008;**283**:33685–97.

7. Ko HL, Zhuo Z, Ren EC. HNF4α combinatorial isoform heterodimers activate distinct gene targets that differ from their corresponding homodimers. *Cell Rep* 2019;**26**:2549–2557.e3.

8. Duda K, Chi YI, Shoelson SE. Structural basis for HNF-4alpha activation by ligand and coactivator binding. *J Biol Chem* 2004;**279**:23311–6.

9. Ryffel GU. Mutations in the human genes encoding the transcription factors of the hepatocyte nuclear factor (HNF)1 and HNF4 families: functional and pathological consequences. *J Mol Endocrinol* 2001;**27**:11–29.

10. Pearson ER, Boj SF, Steele AM. *et al.* Macrosomia and hyperinsulinaemic hypoglycaemia in patients with heterozygous mutations in the HNF4A gene. *PLoS Med* 2007;**4**:e118.

11. Hamilton AJ, Bingham C, McDonald TJ. *et al.* The HNF4A R76W mutation causes atypical dominant Fanconi syndrome in addition to a beta cell phenotype. *J Med Genet* 2014;**51**:165–9.

12. Hansen SK, Rose CS, Glumer C. *et al.* Variation near the hepatocyte nuclear factor (HNF)-4alpha gene associates with type 2 diabetes in the Danish population. *Diabetologia* 2005;**48**:452–8.

13. Mahajan A, Wessel J, Willems SM. *et al.* Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. *Nat Genet* 2018;**50**:559–71.

14. Panel, M.D.V.C.E. 2021. https://clinicalgenome.org/site/assets/files/7133/clingen_diabetes_acmg_specifications_v1_1.pdf, in press.

15. Colclough K, Bellanne-Chantelot C, Saint-Martin C. *et al.* Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 alpha and 4 alpha in maturity-onset diabetes of the young and hyperinsulinemic hypoglycemia. *Hum Mutat* 2013;**34**:669–85.

16. Laver TW, Colclough K, Shepherd M. *et al.* The common p.R114W HNF4A mutation causes a distinct clinical subtype of monogenic diabetes. *Diabetes* 2016;**65**:3212–7.

17. Wright CF, West B, Tuke M. *et al.* Assessing the pathogenicity, penetrance, and expressivity of putative disease-causing variants in a population setting. *Am J Hum Genet* 2019;**104**:275–86.

18. Flannick J, Beer NL, Bick AG. *et al.* Assessing the phenotypic effects in the general population of rare variants in genes for a dominant Mendelian form of diabetes. *Nat Genet* 2013;**45**:1380–5.

19. Malikova J, Kaci A, Dusatkova P. *et al.* Functional analyses of HNF1A-MODY variants refine the interpretation of identified sequence variants. *J Clin Endocrinol Metab* 2020;**105**:e1377–86.

20. Althari S, Najmi LA, Bennett AJ. *et al.* Unsupervised clustering of missense variants in HNF1A using multidimensional

functional data aids clinical interpretation. *Am J Hum Genet* 2020;**107**:670–82.

21. Richards S, Aziz N, Bale S. *et al*. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;**17**:405–24.

22. Cubuk H, Yalcin Capan O. A review of functional characterization of single amino acid change mutations in HNF transcription factors in MODY pathogenesis. *Protein J* 2021;**40**:348–60.

23. Beaudry JB, Pierreux CE, Hayhurst GP. *et al*. Threshold levels of hepatocyte nuclear factor 6 (HNF-6) acting in synergy with HNF-4 and PGC-1alpha are required for time-specific gene expression during liver development. *Mol Cell Biol* 2006;**26**:6037–46.

24. Marchesin V, Perez-Marti A, Le Meur G. *et al*. Molecular basis for autosomal-dominant renal Fanconi syndrome caused by HNF4A. *Cell Rep* 2019;**29**:4407–4421.e5.

25. Chandra V, Huang P, Potluri N. *et al*. Multidomain integration in the structure of the HNF-4α nuclear receptor complex. *Nature* 2013;**495**:394–8.

26. Sun K, Montana V, Chellappa K. *et al*. Phosphorylation of a conserved serine in the deoxyribonucleic acid binding domain of nuclear receptors alters intracellular localization. *Mol Endocrinol* 2007;**21**:1297–311.

27. Yang Q, Yamagata K, Yamamoto K. *et al*. R127W-HNF-4alpha is a loss of function mutation but not a rare polymorphism and causes type II diabetes in a Japanese family with MODY1. *Diabetologia* 2000;**43**:520–4.

28. Najmi LA, Aukrust I, Flannick J. *et al*. Functional investigations of HNF1A identify rare variants as risk factors for type 2 diabetes in the general population. *Diabetes* 2017;**66**:335–46.

29. Zhu Q, Yamagata K, Miura A. *et al*. T130I mutation in HNF-4alpha gene is a loss-of-function mutation in hepatocytes and is associated with late-onset type 2 diabetes mellitus in Japanese subjects. *Diabetologia* 2003;**46**:567–73.

30. Ek J, Rose CS, Jensen DP. *et al*. The functional Thr130Ile and Val255Met polymorphisms of the hepatocyte nuclear factor-4alpha (HNF4A): gene associations with type 2 diabetes or altered beta-cell function among Danes. *J Clin Endocrinol Metab* 2005;**90**:3054–9.

31. Madsen A, Bjune JI, Bjorkhaug L. *et al*. The cAMP-dependent protein kinase downregulates glucose-6-phosphatase expression through RORalpha and SRC-2 coactivator transcriptional activity. *Mol Cell Endocrinol* 2016;**419**:92–101.

32. Garnier J, Osguthorpe DJ, Robson B. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J Mol Biol* 1978;**120**:97–120.

33. Sasaki S, Hara A, Sakaguchi M. *et al*. Hepatocyte nuclear factor 4α regulates megalin expression in proximal tubular cells. *Biochem Biophys Rep* 2019;**17**:87–92.

34. Triplitt CL. Understanding the kidneys' role in blood glucose regulation. *Am J Manag Care* 2012;**18**:S11–6.

35. Gerich JE, Meyer C, Woerle HJ. *et al*. Renal gluconeogenesis: its importance in human glucose homeostasis. *Diabetes Care* 2001;**24**:382–91.

36. Ferrer J. A genetic switch in pancreatic beta-cells: implications for differentiation and haploinsufficiency. *Diabetes* 2002;**51**:2355–62.

37. Keller AD. Model genetic circuits encoding autoregulatory transcription factors. *J Theor Biol* 1995;**172**:169–85.

38. Thomas R, Thieffry D, Kaufman M. Dynamical behaviour of biological regulatory networks—I. Biological role of feedback loops and practical use of the concept of the loop-characteristic state. *Bull Math Biol* 1995;**57**:247–76.

39. Soutoglou E, Katrakili N, Talianidis I. Acetylation regulates transcription factor activity at multiple levels. *Mol Cell* 2000;**5**:745–51.

40. Kaci A, Keindl M, Solheim MH. *et al*. The E3 SUMO ligase PIAS-gamma is a novel interaction partner regulating the activity of diabetes associated hepatocyte nuclear factor-1alpha. *Sci Rep* 2018;**8**:12780.

41. Rodrigues CH, Pires DE, Ascher DB. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Res* 2018;**46**:W350–5.

42. Yin S, Ding F, Dokholyan NV. Eris: an automated estimator of protein stability. *Nat Methods* 2007;**4**:466–7.

43. Gonnelli G, Rooman M, Dehouck Y. Structure-based mutant stability predictions on proteins of unknown structure. *J Biotechnol* 2012;**161**:287–93.

44. Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins* 2006;**62**:1125–32.

45. Parthiban V, Gromiha MM, Schomburg D. CUPSAT: prediction of protein stability upon point mutations. *Nucleic Acids Res* 2006;**34**:W239–42.

46. Pan Y, Zhou S, Guan J. Computationally identifying hot spots in protein-DNA binding interfaces using an ensemble approach. *BMC Bioinformatics* 2020;**21**:384.

47. Peng Y, Sun L, Jia Z. *et al*. Predicting protein-DNA binding free energy change upon missense mutations using modified MM/PBSA approach: SAMPDI webserver. *Bioinformatics* 2018;**34**:779–86.

48. Ioannidis NM, Rothstein JH, Pejaver V. *et al*. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet* 2016;**99**:877–85.