

1
2
3 In silico prediction of the effects of mutations in the human mevalonate
4
5
6 kinase gene: towards a predictive framework for mevalonate kinase
7
8
9 deficiency

10
11
12
13
14
15 Claire Browne¹ and David J. Timson^{1,2*}
16

17
18 ¹ School of Biological Sciences, Queen's University Belfast, Medical Biology Centre, 97 Lisburn Road,
19
20 Belfast, BT9 7BL. UK.
21

22
23 ² Institute for Global Food Security, Queen's University Belfast, 18-30 Malone Road, Belfast, BT9
24
25 5BN. UK.
26

27
28
29
30
31 * Corresponding author.
32

33
34 Address: School of Biological Sciences, Queen's University Belfast, Medical Biology Centre, 97
35
36 Lisburn Road, Belfast, BT9 7BL. UK.
37

38
39 Telephone +44(0)28 9097 5875
40

41
42 Fax +44(0)28 9097 5877
43

44
45 Email d.timson@qub.ac.uk
46
47

48
49
50
51 Running title: **Effects of mutations in human mevalonate kinase**
52
53
54
55
56
57
58
59
60

ABSTRACT

1
2
3
4
5
6 Mevalonate kinase (MVK) catalyses the phosphorylation of mevalonate. Deficiency of MVK is
7
8 associated with two rare periodic fever syndromes, mevalonic aciduria (MA), a severe form and
9
10 hyper-Immunoglobulin-D syndrome (HIDS), a milder form. An in silico approach was used to analyse
11
12 the physicochemical and structural effects of 47 disease-associated variants of MVK. A further 20
13
14 variants, which are present in human genome databases, were also analysed. Variants associated
15
16 with MA are clustered into a “hotspot” consisting of residues 8-35 and 234-338 and tended to result
17
18 in a prediction of severely reduced protein stability. Four of the uncharacterised variants, p.H24P,
19
20 p.G198R, p. R253W and p.G335S were likely to be associated with MA. This method could be used as
21
22 the basis for initial predictions of severity when new MVK variants are discovered.
23
24
25
26
27
28

29 Keywords: periodic fever syndrome; mevalonic aciduria; hyper-Immunoglobulin-D syndrome; HIDS;
30
31 protein stability; disease prediction
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

INTRODUCTION

Mevalonate kinase (MVK) deficiency is associated with a wide range of systemic diseases ranging from the most severe form, mevalonic aciduria (MA; OMIM#610377) to a milder form, hyper-immunoglobulin D syndrome (HIDS; OMIM#60920) (Drenth et al. 1999, Haas & Hoffmann 2006, Houten et al. 2000b). These rare autosomal recessive disorders are caused by a mutation in the MVK gene, which is found on the long arm of chromosome 12 (12q24) (Haas & Hoffmann 2006, Houten et al. 2000b, Houten et al. 2000a). Mevalonate kinase (EC 2.7.1.36) is an essential enzyme in the cholesterol pathway (Buhaescu & Izzedine 2007). It is a dimeric protein that catalyses the ATP-dependent phosphorylation of mevalonate to 5-phosphomevalonate (Potter & Miziorko 1997). This step follows the one catalysed by the highly regulated enzyme and statin target, 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) reductase (EC 1.1.1.34). Although there is no effective treatment for MVK deficiency, preliminary evidence has shown that Simvastatin, an HMG-CoA reductase inhibitor, may be useful in treating inflammatory attacks in HIDS (Simon et al. 2004).

MVK is part of the GHMP family of kinases (Bork et al. 1993, Timson 2007). Each subunit has two domains with the active site located in a cleft at the domain interface (Yang et al. 2002, Fu et al. 2002, Sgraja et al. 2007, Fu et al. 2008). **The enzyme is dimeric.** Thus, mutations causing amino acid changes at the dimer interface may disrupt dimer formation leading to an inactive or unstable protein (Fu et al. 2002). The crystal structure of rat mevalonate kinase in complex with MgATP revealed that an Mg^{2+} ion is co-ordinated by both β - and γ -phosphates of ATP and side chains of Glu-193 and Ser-146 (Fu et al. 2002). Asp204 was found to make a salt bridge with Lys-13, which interacts with the γ -phosphate (Fu et al. 2002). Lys-13 affects the pK_a of the C5 hydroxyl of mevalonate, while Asp-204 abstracts the proton from this hydroxyl. The resulting penta-coordinated γ -phosphoryl group may be stabilized by Mg^{2+} , Lys-13, and Glu-193 (Fu et al. 2002). **The structure of the human MVK enzyme is very similar to the rat one (rmsd 0.81 Å) and the mechanism of reaction is expected to be identical with Asp-204 acting as a base in the active site (Fu et al. 2008).**

1
2
3 MA and HIDS are rare autosomal recessive disorders characterised by recurrent periodic fevers and
4
5 generalised inflammation (Haas & Hoffmann 2006, Frenkel et al. 2000). Typical symptoms of the
6
7 severe form, MA, can include psychomotor retardation, ataxia, failure to thrive, dysmorphic features
8
9 with recurrent episodes of fevers, lymphadenopathy and rashes (Frenkel et al. 2000, van der Burgh
10
11 et al. 2013). These symptoms typically present in infancy and patients usually die in childhood
12
13 (Frenkel et al. 2000). Suffers of MA can be confirmed based on their biochemical, clinical and genetic
14
15 data (Drenth et al. 1999). The residual activity of MVK in MA sufferers is below 0.5%, leading to a
16
17 characteristic build-up of mevalonic acid (Prietsch et al. 2003).

18
19
20
21 HIDS is a milder phenotype of MVK deficiency with more documented cases (Haas & Hoffmann
22
23 2006, Frenkel et al. 2000, van der Burgh et al. 2013, Grose 2005). Symptoms usually present in the
24
25 first year of life and include recurrent lifelong episodes of fever, rashes, abdominal pain and
26
27 lymphadenopathy (Haas & Hoffmann 2006, Frenkel et al. 2000, McDermott & Frenkel 2001, Haas et
28
29 al. 2001, Hoffman et al. 2001). A high level of IgD and IgA accompanies these symptoms and is often
30
31 used as a diagnostic confirmation of HIDS (Drenth et al. 1999, Houten et al. 2000b, Buhaescu &
32
33 Izzedine 2007). Moreover the residual activity of MVK in HIDS is around 1-20% (Houten et al. 2000b).
34
35 HIDS can be distinguished from MA by the lack of abnormal levels of mevalonic acid present as well
36
37 as the lack of neurological manifestations which feature in MA (Hoffman et al. 2001). Furthermore
38
39 unlike individuals with MA, people suffering with HIDS do not experience any symptoms between
40
41 fever episodes and typically have a normal life expectancy (Haas & Hoffmann 2006).
42
43
44

45
46 The most frequent mutation associated with HIDS is c.1129G>A (rs28934897) resulting in an amino
47
48 acid change from valine to isoleucine at position 377 (p.V377I) (Cuisset et al. 2001). This is found in
49
50 approximately 80% of HIDS sufferers and has never been reported in MA (Cuisset et al. 2001). MVK
51
52 deficiency is particularly common in northern Europe (especially the The Netherlands and France)
53
54 but cases have been reported worldwide (Prietsch et al. 2003). Approximately 80 mutations have
55
56 been reported to cause mevalonate kinase deficiency (Haas & Hoffmann 2006).
57
58
59
60

1
2
3 Overlapping phenotypes of MA and HIDS have been observed in an intermediate form, although few
4
5 cases have been documented (Hoffman et al. 2001). However it is important to note that the
6
7 relationship between mutation and reported phenotype may not always be straightforward.

8
9 Phenotypes are classified based on individual clinical observation and can be influenced by
10
11 environmental factors; for example, the symptoms of MVK deficiency appear to be amplified by
12
13 increased temperature (Houten et al. 2002, Tricarico et al. 2013).

14
15
16
17 It is not known precisely how each disease-associated mutation alters the enzyme structure or how
18
19 this contributes to the different levels of severity or pathology. It is assumed that the residual
20
21 activity of MVK determines whether or not a patient has MA or HIDS, therefore a mutation
22
23 associated with the severe form must disrupt the enzyme's structure and function significantly to
24
25 decrease its activity below 0.5%. At the biochemical level this is likely to result from effects on
26
27 protein folding and stability. A bioinformatics approach was taken to understand the functional and
28
29 structural contribution of molecular alterations in MVK and how these correlate to the associated
30
31 severity. This was then applied to variants of MVK whose disease association is currently unknown.

32
33 The approaches can be broadly arranged into three approaches: sequence and evolutionary
34
35 conservation based methods, protein sequence and structure-based methods, and supervised
36
37 learning methods. Similar approaches have been successful in the investigation of type I
38
39 galactosemia (galactose 1-phosphate uridylyltransferase deficiency), type III galactosemia (UDP-
40
41 galactose 4'-epimerase deficiency), hyperargininemia (arginase 1 deficiency) and apparent
42
43 mineralocorticoid excess (11 β -hydroxysteroid dehydrogenase type 2 deficiency) to provide the basis
44
45 for predicting the severity of newly discovered mutations (Carvalho et al. 2012, Manning et al. 2010,
46
47 Facchiano & Marabotti 2010, d'Acierno et al. 2009, d'Acierno et al. 2014, McCorvie & Timson 2013).

48
49
50
51
52
53
54
55 MATERIALS AND METHODS
56
57
58
59
60

1
2
3 Datasets, sequence information and data analysis
4

5
6 Literature searches were used to identify disease-associated variants. Where available, the clinical
7
8 symptoms, observed residual activity of the MVK variant, any biochemical analysis (including any
9
10 experimental conditions), environmental factors and the genetic background were noted.

11
12 Characterised variants and uncharacterised missense mutations from exome sequencing were
13
14 identified from the databases: NCBI dbSNP (<http://www.ncbi.nlm.nih.gov/SNP/>), UniProt
15
16 (<http://www.uniprot.org>), and Infevers (<http://fmf.igh.cnrs.fr/ISSAID/infevers/>). Excel (Microsoft)
17
18 and GraphPad Prism 5.0 (GraphPad Software) were used to record scores, perform statistical
19
20 analysis and obtain graphical representations of the results.
21
22

23
24
25
26
27 Secondary and tertiary structural investigations

28
29 Structural studies were based on the crystal structure of human MVK (PDB: 2R3V) (Fu et al. 2008).
30
31 Structures were visualised using PyMol (<http://www.pymol.com>) and were computationally solvated
32
33 and energy minimized using YASARA (<http://www.yasara.org>) (Krieger et al. 2009). Sequence
34
35 variations were introduced in silico using the Mutate function in PyMol and these variant forms of
36
37 the protein were minimised using YASARA. Thus, all structural comparisons were made using
38
39 minimised structures. Each variant's structure was further analysed using LS-SNP/PDB ([http://ls-](http://ls-snp.icm.jhu.edu/ls-snp-pdb/main)
40
41 [snp.icm.jhu.edu/ls-snp-pdb/main](http://ls-snp.icm.jhu.edu/ls-snp-pdb/main)). This tool provided solvent accessibility scores, variant's position
42
43 at domain interface, location in the 3D structure, if it is exposed or buried, the secondary structure
44
45 where the variant is found and a 3D representation of the variant in the structure. The GETAREA
46
47 server (<http://curie.utmb.edu/getarea.html>) was used to determine the change in surface and
48
49 buried atoms from the variant compared to the wild type (Fraczkiewicz & Braun 1998).
50
51
52
53
54
55
56
57
58
59
60

1
2
3 Prediction of the physicochemical effect of each missense mutation on MVK structure, stability and
4
5 function

6
7
8 Bioinformatics tools were used to investigate a range of physicochemical effects of each variant on
9
10 MVK's structure and function. The following tools were used to estimate the change in protein
11
12 stability: I-Mutant 3.0 Binary and Ternary classification

13
14 (<http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/I-Mutant3.0.cgi>) (Capriotti et al. 2005,
15
16 Capriotti et al. 2006); mCSM (<http://bleoberis.bioc.cam.ac.uk/mcsm/>) (Pires et al. 2014); SDM score
17
18 (<http://mordred.bioc.cam.ac.uk/~sdm/sdm.php>) (Worth et al. 2011, Worth et al. 2007); Mupro
19
20 (<http://mupro.proteomics.ics.uci.edu/>) (Cheng et al. 2006); iStable score
21
22 (<http://predictor.nchu.edu.tw/iStable/>) (Chen et al. 2013); PredictSNP 1.0
23
24 (<http://loschmidt.chemi.muni.cz/predictsnp/>) (Bendl et al. 2014); Meta-SNP
25
26 (<http://snps.biofold.org/meta-snp/>) (Capriotti et al. 2013); KD4V score
27
28 (<http://decryphon.igbmc.fr/kd4v>) (Luu et al. 2012).

29
30
31
32 The tools used to estimate the change in free energy resulting from the point mutations (all
33
34 calculated at pH 7.5, 30 °C) were: I-Mutant 3.0 DDG, SVM3 and SVM2 values (Capriotti et al. 2006);
35
36 Fold-X DDG value (<http://foldx.crg.es>) (Schymkowitz et al. 2005); **PoPMuSiC 2.1**
37
38 (<http://dezyme.com/>) (Dehouck et al. 2011); CUPSAT DDG (<http://cupsat.tu-bs.de>) (Parthiban et al.
39
40 2006, Parthiban et al. 2007); iStable DDG (Chen et al. 2013); GETAREA energy change
41
42 (<http://curie.utmb.edu/getarea.html>) (Fraczkiewicz & Braun 1998).

43
44
45
46 The tools used to predict change in binding affinity were: mCSM Protein-Protein affinity prediction
47
48 (Pires et al. 2014); BeAtMuSiC binding affinity prediction and score
49
50 (<http://babylone.ulb.ac.be/beatmusic>) (Dehouck et al. 2013). The tools used predict the change in
51
52 solvent accessibility were: **PoPMuSiC 2.1** (Dehouck et al. 2011); CUPSAT RSA (<http://cupsat.tu->
53
54 bs.de) (Parthiban et al. 2007); SDM score (Worth et al. 2011, Worth et al. 2007); BeAtMuSiC
55
56 (Dehouck et al. 2013).

1
2
3 In addition, further physicochemical analysis was carried out using SNP effect 4.0 server
4
5 (<http://snpeffect.switchlab.org>) (De Baets et al. 2012) which uses four programs to predict the
6
7 biochemical effects of missense mutations; TANGO (<http://tango.crg.es/>) (Fernandez-Escamilla et al.
8
9 2004) which predicts the tendency to aggregate; Fold-X (Schymkowitz et al. 2005) which predicts
10
11 changes in stability, WALTZ (<http://www.switchlab.org/bioinformatics/waltz>) (Maurer-Stroh et al.
12
13 2010) which predicts the tendency to form amyloids and LIMBO
14
15 (<http://www.switchlab.org/bioinformatics/limbo>) (Van Durme et al. 2009) which predicts the ability
16
17 to bind chaperones. A estimation of the change in hydrogen bond satisfaction was also carried out
18
19 using SDM (Worth et al. 2011, Worth et al. 2007). The KD4V server (Luu et al. 2012) provided
20
21 additional scores in the change in size, charge polarity, modifications and hydrophobicity. Change in
22
23 hydrophobicity scores were calculated using data from (Black & Mould 1991).
24
25
26
27
28
29

30 Multiple sequence alignment and evolutionary conservation based methods

31
32
33 Multiple sequence alignment was carried out using Clustal Omega
34
35 (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) (Sievers et al. 2011) from sequences gathered from
36
37 UniProt database (<http://www.uniprot.org>). Only those classified as MVK and with “reviewed”
38
39 status were used. A total of 19 sequences were obtained, which included one from yeast, four from
40
41 vertebrates, one from slime mould, one from plants, one from Archaea and eleven from eubacteria.
42
43
44 The resulting alignment (**Supplementary Figure S1**) was used with the Scorecons server
45
46 (https://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/valdar/scorecons_server.pl) based on the
47
48 Valdar01 score (Valdar 2002). The alignment was also used to calculate the degree of tolerance of
49
50 each amino acid alteration using the SIFT server (<http://sift.jcvi.org/>) (Kumar et al. 2009) and the
51
52 PROVEAN server (<http://provean.jcvi.org/index.php>) (Choi et al. 2012), both of which are based on
53
54 updated versions of Ensembl gene annotation (GRCh37 Ensembl 66) and NCBI dbSNP database
55
56 (Build 137).
57
58
59
60

1
2
3 Additional investigations of the tolerance to change at positions which are altered in the disease-
4 associated variants was carried out using the tools: LS-SNP/PDB (<http://ls-snp.icm.jhu.edu/ls-snp->
5 [pdb/](http://ls-snp.icm.jhu.edu/ls-snp-pdb/)) (Ryan et al. 2009); SNPs&GO (<http://snps.biofold.org/snps-and-go/pages/help.html>)
6
7 (Calabrese et al. 2009); PhD-SNP (<http://snps.biofold.org/phd-snp/phd-snp.html>) (Capriotti et al.
8
9 2006); PANTHER (<http://www.pantherdb.org/tools/csnpScoreForm.jsp>) (Brunham et al. 2005);
10
11 GenMAPP (<http://www.genmapp.org>) (Salomonis et al. 2007); PolyPhen 2
12
13 (<http://genetics.bwh.harvard.edu/pph2/>) (Adzhubei et al. 2010, Adzhubei et al. 2013); nsSNP
14
15 Analyzer (<http://snpanalyzer.uthsc.edu>) (Bao et al. 2005); FI mutation assessor
16
17 (<http://mutationassessor.org/v1>) (Reva et al. 2011); KD4V (Luu et al. 2012); YALE MU2A
18
19 (<http://krauthammerlab.med.yale.edu/mu2a>) (Garla et al. 2011).
20
21
22
23
24
25
26
27

28 Supervised learning methods for combined overall predictions of each variant severity
29
30

31 Supervised learning methods that include a variety of bioinformatics tools to compare variants and
32
33 provide an overall prediction on the effect of the variant on the protein's structure, function and
34
35 pathology were used. Programs that provided this overall score, predicting if the mutation was
36
37 disease causing, damaging, destabilizing or deleterious were: I-Mutant 3.0 Prediction of Disease
38
39 (Capriotti et al. 2006); SIFT (Kumar et al. 2009); PROVEAN prediction (Choi et al. 2012); PolyPhen 2
40
41 prediction (Adzhubei et al. 2013); mCSM prediction (Pires et al. 2014); SNAP prediction
42
43 (<https://www.rostlab.org/services/snap/>) (Bromberg & Rost 2007); PANTHER (Brunham et al. 2005);
44
45 PhD-SNP predictor (Capriotti et al. 2006); nsSNP Analyzer (Bao et al. 2005); Mupro prediction (Cheng
46
47 et al. 2006); **PopMuSiC 2.1 (Dehouck et al. 2011)**; CUPSAT (Parthiban et al. 2007); iStable (Chen et al.
48
49 2013); PredictSNP (Bendl et al. 2014); MAPP (Salomonis et al. 2007); Meta-SNP (Capriotti et al.
50
51 2013); FI mutation assessor (Reva et al. 2011); BeAtMuSiC (Dehouck et al. 2013).
52
53
54
55
56
57
58
59
60

RESULTS AND DISCUSSION

MVK alterations associated with MA tend to be located in the central cleft, near the active site

Literature and database searches identified forty-seven MVK mutations for which there was sufficient biochemical and clinical information to be included in this study. Four of these were associated with the intermediate severity MVK deficiency, twelve with the severe form (MA) and the remainder with the mild form (HIDS). The location of the changes in the protein sequence showed that there was a relationship between the position of the variant and associated severity, with severe and intermediate mutations only occurring between amino acid positions 8-35 and 234-338 (Figure 1). These two sequences are located around the inside of the protein's cleft close to the active site and around the domain interface. Four of the variations associated with the severe form (p.L264F, p.L265P, p.I268H and p.N301T) occur in α -helices at the dimer interface. To our knowledge there is no published information on the effect of these (or other) mutations on the ability of MVK to dimerise; nor is there any information on whether, or not, dimerization is required for full activity of human MVK. However, MVK from some bacteria is monomeric, suggesting that the oligomeric state may have limited impact on activity (Hedl & Rodwell 2004, Voynova et al. 2004). The alterations associated with HIDS were more widespread throughout the MVK structure (Figure 1). The position of the variants in the 3D structure from LS-SNP/PDB showed that the majority of variants associated with severe and intermediate forms of the disease affected buried residues, while the variants associated with HIDS affected a mixture of buried, intermediate and exposed residues.

Computational analysis tends to place the MVK variants into two groups

All the results from the physicochemical analysis of the disease-associated and uncharacterised variants are given in Supplementary Table S1. There was a tendency to predict decreasing structural

1
2
3 stability with increasing severity of the disease-associated variants. The programs that gave this
4 trend with were: I-Mutant DDG, solvent accessibility, FOLD-X, mCSM stability, Mupro, PoPMuSiC 2.1,
5 BeAtMuSiC, CUPSAT (RSA), iStable DDG, GETAREA, SDM, BETAMUSIC, and KD4V. The change in
6 hydrophobicity tended to increase with severity (data not shown). In the majority of cases, however,
7 there was no statistical significance ($p>0.05$) when the three groups (mild, intermediate and severe)
8 were compared. In general, alterations at more highly conserved residues were more likely to result
9 in severe forms of MVK deficiency. This result was supported by predictions from: POLYPHEN-2,
10 SAM, SCORECONS, PHD-SNP, SNP&GO, PANTHER, MAPP, FI-mutant and 4D4V WT residue
11 representation. However, like the physiochemical scores, there was no significant difference
12 between the scores when comparing the results from the three groups with the majority of the
13 programs (data not shown).
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31

Supervised learning methods predict effects on MVK stability using a combination of methods

32
33 Programs that used supervised learning methods for an overall prediction were used to predict each
34 variant's effect on MVK's stability. It should be noted that this approach is based on the assumption
35 that all the prediction tools are equally good. However, the percentage score provides an overall
36 consensus which has been found to be more accurate than only using one or two tools (Zhao et al,
37 2014). The results in Supplementary Table S2 are based on 28 prediction tools and show that, in
38 general, the percentage of tools predicting a reduction in MVK stability slightly increases with
39 increasing severity. For each group the percentage predicting a decrease ranged from: 28.6-89.3
40 (mild), 53.6-89.3 (intermediate) and 64.3-92.9 (severe). Although there is an increasing trend, the
41 discrimination between groups was not sufficient to make meaningful predictions about variants of
42 unknown severity.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 The intermediate form was not well detected by these methods
4

5
6 Across the three different prediction methods, the intermediate group was not well discriminated.
7

8 In general it showed results similar to, or greater than, the severe group (Supplementary Table S1).
9

10 This suggests that either these methods are not sufficiently discriminating or that the intermediate
11 group would be better considered as part of the severe group. Combining these two groups enabled
12 discrimination between the combined intermediate/severe group and the mild group on the basis of
13 the physicochemical scores (Figure 2). Therefore, in developing a predictive framework, these two
14 groups were considered together.
15
16
17
18
19

20
21
22
23
24
25 Towards a predictive framework: four key postulates
26

27
28 Based on the data (summarised in Supplementary Table S2), four postulates were made:
29

- 30
31 1. The decrease in enzyme activity is correlated with severity of disease. In the majority of
32 cases, this decrease results from improper folding of the protein.
33
34
35
36 2. Changes to highly conserved residues are more likely to result in improper folding.
37
38
39 3. Intermediate mutations should be treated with the severe group due to their similarity to
40 this group.
41
42
43
44 4. For a severe or intermediate phenotype to occur, the mutation is likely to cause a change in
45 the "hotspot" region, identified from structural analysis, i.e. around residues 8-35 and 234-
46 338 in the protein sequence.
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Prediction of the likely clinical severity associated with uncharacterised variants

There are a number of variants which have been identified through genome and exome sequencing projects, but have not yet been characterised in terms of their associated disease severity. Applying the postulates above, four of these variants are predicted to be associated with the intermediate or severe form (MA). These are p.H24P, p.G198R, p.R253W and p.G335S (Supplementary Table S3).

Three of these uncharacterised variants were predicted to cause either the mild form (HIDS) or no disease and the rest were predicted to be associated with HIDS (Supplementary Table S3).

Experimental testing will be necessary to confirm these predictions. Molecular dynamics simulations may be valuable in explaining why disease-associated variants affect protein stability and would also be useful in investigating the properties of these uncharacterised variants. It is further hypothesised that the MVK variants which are predicted to be associated with MA will have very low activities and stabilities in vitro. It is also expected that cells homozygous for the corresponding mutations are likely to be deficient in the mevalonate pathway and will accumulate mevalonate.

Conclusions

The analyses presented here suggest that point mutations in the MVK gene which result in changes to the protein coding sequence subtly alter the structure and stability of MVK, thus affecting its activity. These effects are greater when the change occurs in a well-conserved residue. Variants associated with MA generally result in changes to the highly conserved "hotspot" region and also result in greater changes to stability and other biophysical parameters. In addition to providing a predictive framework, these results also suggest that pharmacological interventions to stabilise the MVK protein ("small molecule chaperone therapy") may be possible in the treatment of MA and the alleviation of the feverish episodes in HIDS.

ACKNOWLEDGEMENTS

We thank Dr Thomas J. McCorvie (Structural Genomics Centre, University of Oxford, UK) for helpful discussions. This work received no funding from any external source.

The authors have no conflicts of interest to declare.

For Peer Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 References
4
5

6 Adzhubei, I., Jordan, D.M. & Sunyaev, S.R. (2013) Predicting Functional Effect of Human Missense
7 Mutations Using PolyPhen-2. *Curr.Protoc.Hum.Genet.* Chapter 7, Unit7.20.

8 Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S. &
9 Sunyaev, S.R. (2010) A method and server for predicting damaging missense mutations.
10 *Nat.Methods.* 7, 248-249.

11 Bao, L., Zhou, M. & Cui, Y. (2005) nsSNPAnalyzer: identifying disease-associated nonsynonymous
12 single nucleotide polymorphisms. *Nucleic Acids Res.* 33, W480-2.

13 Bendl, J., Stourac, J., Salanda, O., Pavelka, A., Wieben, E.D., Zendulka, J., Brezovsky, J. & Damborsky,
14 J. (2014) PredictSNP: robust and accurate consensus classifier for prediction of disease-related
15 mutations. *PLoS Comput.Biol.* 10, e1003440.

16 Black, S.D. & Mould, D.R. (1991) Development of hydrophobicity parameters to analyze proteins
17 which bear post- or cotranslational modifications. *Anal.Biochem.* 193, 72-82.

18 Bork, P., Sander, C. & Valencia, A. (1993) Convergent evolution of similar enzymatic function on
19 different protein folds: the hexokinase, ribokinase, and galactokinase families of sugar kinases.
20 *Protein Sci.* 2, 31-40.

21 Bromberg, Y. & Rost, B. (2007) SNAP: predict effect of non-synonymous polymorphisms on function.
22 *Nucleic Acids Res.* 35, 3823-3835.

23 Brunham, L.R., Singaraja, R.R., Pape, T.D., Kejariwal, A., Thomas, P.D. & Hayden, M.R. (2005)
24 Accurate prediction of the functional significance of single nucleotide polymorphisms and mutations
25 in the ABCA1 gene. *PLoS Genet.* 1, e83.

26 Buhaescu, I. & Izzedine, H. (2007) Mevalonate pathway: a review of clinical and therapeutical
27 implications. *Clin.Biochem.* 40, 575-584.

28 Calabrese, R., Capriotti, E., Fariselli, P., Martelli, P.L. & Casadio, R. (2009) Functional annotations
29 improve the predictive score of human disease-related mutations in proteins. *Hum.Mutat.* 30, 1237-
30 1244.

31 Capriotti, E., Altman, R.B. & Bromberg, Y. (2013) Collective judgment predicts disease-associated
32 single nucleotide variants. *BMC Genomics.* 14 Suppl 3, S2-2164-14-S3-S2. Epub 2013 May 28.

33 Capriotti, E., Calabrese, R. & Casadio, R. (2006) Predicting the insurgence of human genetic diseases
34 associated to single point protein mutations with support vector machines and evolutionary
35 information. *Bioinformatics.* 22, 2729-2734.

36 Capriotti, E., Fariselli, P. & Casadio, R. (2005) I-Mutant2.0: predicting stability changes upon mutation
37 from the protein sequence or structure. *Nucleic Acids Res.* 33, W306-10.

38 Carvalho, D.R., Brand, G.D., Brum, J.M., Takata, R.I., Speck-Martins, C.E. & Pratesi, R. (2012) Analysis
39 of novel ARG1 mutations causing hyperargininemia and correlation with arginase I activity in
40 erythrocytes. *Gene.* 509, 124-130.

41 Chen, C.W., Lin, J. & Chu, Y.W. (2013) iStable: off-the-shelf predictor integration for predicting
42 protein stability changes. *BMC Bioinformatics.* 14 Suppl 2, S5-2105-14-S2-S5. Epub 2013 Jan 21.
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 Cheng, J., Randall, A. & Baldi, P. (2006) Prediction of protein stability changes for single-site
4 mutations using support vector machines. *Proteins*. 62, 1125-1132.
5
6 Choi, Y., Sims, G.E., Murphy, S., Miller, J.R. & Chan, A.P. (2012) Predicting the functional effect of
7 amino acid substitutions and indels. *PLoS One*. 7, e46688.
8
9 Cuisset, L., Drenth, J.P., Simon, A., Vincent, M.F., van der Velde Visser, S., van der Meer, J.W.,
10 Grateau, G., Delpech, M. & International Hyper-IgD Study Group. (2001) Molecular analysis of MVK
11 mutations and enzymatic activity in hyper-IgD and periodic fever syndrome. *Eur.J.Hum.Genet.* 9,
12 260-266.
13
14 d'Acerno, A., Facchiano, A. & Marabotti, A. (2014) GALT protein database: querying structural and
15 functional features of GALT enzyme. *Hum.Mutat.* 35, 1060-1067.
16
17 d'Acerno, A., Facchiano, A. & Marabotti, A. (2009) GALT protein database, a bioinformatics resource
18 for the management and analysis of structural features of a galactosemia-related protein and its
19 mutants. *Genomics Proteomics Bioinformatics.* 7, 71-76.
20
21 De Baets, G., Van Durme, J., Reumers, J., Maurer-Stroh, S., Vanhee, P., Dopazo, J., Schymkowitz, J. &
22 Rousseau, F. (2012) SNPeff 4.0: on-line prediction of molecular and structural effects of protein-
23 coding variants. *Nucleic Acids Res.* 40, D935-D939.
24
25 Dehouck, Y., Kwasigroch, J.M., Rooman, M. & Gilis, D. (2013) BeAtMuSiC: Prediction of changes in
26 protein-protein binding affinity on mutations. *Nucleic Acids Res.* 41, W333-9.
27
28 Dehouck, Y., Kwasigroch, J.M., Gilis, D. & Rooman, M. (2011) PoPMuSiC 2.1: a web server for the
29 estimation of protein stability changes upon mutation and sequence optimality. *BMC Bioinformatics.*
30 12, 151-2105-12-151.
31
32 Drenth, J.P., Cuisset, L., Grateau, G., Vasseur, C., van de Velde-Visser, S.D., de Jong, J.G., Beckmann,
33 J.S., van der Meer, J.W. & Delpech, M. (1999) Mutations in the gene encoding mevalonate kinase
34 cause hyper-IgD and periodic fever syndrome. *International Hyper-IgD Study Group. Nat.Genet.* 22,
35 178-181.
36
37 Facchiano, A. & Marabotti, A. (2010) Analysis of galactosemia-linked mutations of GALT enzyme
38 using a computational biology approach. *Protein Eng.Des.Sel.* 23, 103-113.
39
40 Fernandez-Escamilla, A.M., Rousseau, F., Schymkowitz, J. & Serrano, L. (2004) Prediction of
41 sequence-dependent and mutational effects on the aggregation of peptides and proteins.
42 *Nat.Biotechnol.* 22, 1302-1306.
43
44 Fraczek, R. & Braun, W. (1998) Exact and efficient analytical calculation of the accessible surface
45 areas and their gradients for macromolecules. *J.Comput.Chem.* 19, 319-333.
46
47 Frenkel, J., Houten, S.M., Waterham, H.R., Wanders, R.J., Rijkers, G.T., Kimpen, J.L., Duran, R., Poll-
48 The, B.T. & Kuis, W. (2000) Mevalonate kinase deficiency and Dutch type periodic fever.
49 *Clin.Exp.Rheumatol.* 18, 525-532.
50
51 Fu, Z., Voynova, N.E., Herdendorf, T.J., Miziorko, H.M. & Kim, J.J. (2008) Biochemical and structural
52 basis for feedback inhibition of mevalonate kinase and isoprenoid metabolism. *Biochemistry.* 47,
53 3715-3724.
54
55 Fu, Z., Wang, M., Potter, D., Miziorko, H.M. & Kim, J.J. (2002) The structure of a binary complex
56 between a mammalian mevalonate kinase and ATP: insights into the reaction mechanism and
57 human inherited disease. *J.Biol.Chem.* 277, 18134-18142.
58
59
60

- 1
2
3 Garla, V., Kong, Y., Szpakowski, S. & Krauthammer, M. (2011) MU2A--reconciling the genome and
4 transcriptome to determine the effects of base substitutions. *Bioinformatics*. 27, 416-418.
5
6 Grose, C. (2005) Periodic fever in children with hyperimmunoglobulinemia D and mevalonate kinase
7 mutations. *Pediatr.Infect.Dis.J.* 24, 573-574.
8
9 Haas, D. & Hoffmann, G.F. (2006) Mevalonate kinase deficiencies: from mevalonic aciduria to
10 hyperimmunoglobulinemia D syndrome. *Orphanet J.Rare Dis.* 1, 13.
11
12 Haas, D., Kelley, R.I. & Hoffmann, G.F. (2001) Inherited disorders of cholesterol biosynthesis.
13 *Neuropediatrics*. 32, 113-122.
14
15 **Hedl, M. & Rodwell, V.W. (2004) Enterococcus faecalis mevalonate kinase. *Protein Sci.* 13, 687-693.**
16
17 Hoffman, H.M., Wanderer, A.A. & Broide, D.H. (2001) Familial cold autoinflammatory syndrome:
18 phenotype and genotype of an autosomal dominant periodic fever. *J.Allergy Clin.Immunol.* 108, 615-
19 620.
20
21 Houten, S.M., Frenkel, J., Rijkers, G.T., Wanders, R.J., Kuis, W. & Waterham, H.R. (2002) Temperature
22 dependence of mutant mevalonate kinase activity as a pathogenic factor in hyper-IgD and periodic
23 fever syndrome. *Hum.Mol.Genet.* 11, 3115-3124.
24
25 Houten, S.M., Frenkel, J., Kuis, W., Wanders, R.J., Poll-The, B.T. & Waterham, H.R. (2000a) Molecular
26 basis of classical mevalonic aciduria and the hyperimmunoglobulinaemia D and periodic fever
27 syndrome: high frequency of 3 mutations in the mevalonate kinase gene. *J.Inherit.Metab.Dis.* 23,
28 367-370.
29
30 Houten, S.M., Wanders, R.J. & Waterham, H.R. (2000b) Biochemical and genetic aspects of
31 mevalonate kinase and its deficiency. *Biochim.Biophys.Acta.* 1529, 19-32.
32
33 Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D. & Karplus, K. (2009)
34 Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four
35 approaches that performed well in CASP8. *Proteins*. 77 Suppl 9, 114-122.
36
37 Kumar, P., Henikoff, S. & Ng, P.C. (2009) Predicting the effects of coding non-synonymous variants
38 on protein function using the SIFT algorithm. *Nat.Protoc.* 4, 1073-1081.
39
40 Luu, T.D., Rusu, A., Walter, V., Linard, B., Poidevin, L., Ripp, R., Moulinier, L., Muller, J., Raffelsberger,
41 W., Wicker, N., Lecompte, O., Thompson, J.D., Poch, O. & Nguyen, H. (2012) KD4v: Comprehensive
42 Knowledge Discovery System for Missense Variant. *Nucleic Acids Res.* 40, W71-5.
43
44 Manning, J.R., Bailey, M.A., Soares, D.C., Dunbar, D.R. & Mullins, J.J. (2010) In silico structure-
45 function analysis of pathological variation in the HSD11B2 gene sequence. *Physiol.Genomics.* 42,
46 319-330.
47
48 Maurer-Stroh, S., Debulpaep, M., Kuemmerer, N., Lopez de la Paz, M., Martins, I.C., Reumers, J.,
49 Morris, K.L., Copland, A., Serpell, L., Serrano, L., Schymkowitz, J.W. & Rousseau, F. (2010) Exploring
50 the sequence determinants of amyloid structure using position-specific scoring matrices.
51 *Nat.Methods.* 7, 237-242.
52
53 McCorvie, T.J. & Timson, D.J. (2013) In silico prediction of the effects of mutations in the human
54 UDP-galactose 4'-epimerase gene: Towards a predictive framework for type III galactosemia. *Gene.*
55 524, 95-104.
56
57 McDermott, M.F. & Frenkel, J. (2001) Hereditary periodic fever syndromes. *Neth.J.Med.* 59, 118-125.
58
59
60

- 1
2
3 Parthiban, V., Gromiha, M.M., Abhinandan, M. & Schomburg, D. (2007) Computational modeling of
4 protein mutant stability: analysis and optimization of statistical potentials and structural features
5 reveal insights into prediction model development. *BMC Struct.Biol.* 7, 54.
6
7 Parthiban, V., Gromiha, M.M. & Schomburg, D. (2006) CUPSAT: prediction of protein stability upon
8 point mutations. *Nucleic Acids Res.* 34, W239-42.
9
10 Pires, D.E., Ascher, D.B. & Blundell, T.L. (2014) mCSM: predicting the effects of mutations in proteins
11 using graph-based signatures. *Bioinformatics.* 30, 335-342.
12
13 Potter, D. & Mizioro, H.M. (1997) Identification of catalytic residues in human mevalonate kinase.
14 *J.Biol.Chem.* 272, 25449-25454.
15
16 Prietsch, V., Mayatepek, E., Krastel, H., Haas, D., Zundel, D., Waterham, H.R., Wanders, R.J., Gibson,
17 K.M. & Hoffmann, G.F. (2003) Mevalonate kinase deficiency: enlarging the clinical and biochemical
18 spectrum. *Pediatrics.* 111, 258-261.
19
20 Reva, B., Antipin, Y. & Sander, C. (2011) Predicting the functional impact of protein mutations:
21 application to cancer genomics. *Nucleic Acids Res.* 39, e118.
22
23 Ryan, M., Diekhans, M., Lien, S., Liu, Y. & Karchin, R. (2009) LS-SNP/PDB: annotated non-synonymous
24 SNPs mapped to Protein Data Bank structures. *Bioinformatics.* 25, 1431-1432.
25
26 Salomonis, N., Hanspers, K., Zambon, A.C., Vranizan, K., Lawlor, S.C., Dahlquist, K.D., Doniger, S.W.,
27 Stuart, J., Conklin, B.R. & Pico, A.R. (2007) GenMAPP 2: new features and resources for pathway
28 analysis. *BMC Bioinformatics.* 8, 217.
29
30 Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F. & Serrano, L. (2005) The FoldX web server:
31 an online force field. *Nucleic Acids Res.* 33, W382-8.
32
33 Sgraja, T., Smith, T.K. & Hunter, W.N. (2007) Structure, substrate recognition and reactivity of
34 *Leishmania major* mevalonate kinase. *BMC Struct.Biol.* 7, 20.
35
36 Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert,
37 M., Soding, J., Thompson, J.D. & Higgins, D.G. (2011) Fast, scalable generation of high-quality protein
38 multiple sequence alignments using Clustal Omega. *Mol.Syst.Biol.* 7, 539.
39
40 Simon, A., Drewe, E., van der Meer, J.W., Powell, R.J., Kelley, R.I., Stalenhoef, A.F. & Drenth, J.P.
41 (2004) Simvastatin treatment for inflammatory attacks of the hyperimmunoglobulinemia D and
42 periodic fever syndrome. *Clin.Pharmacol.Ther.* 75, 476-483.
43
44 Timson, D.J. (2007) GHMP kinases - structures, mechanisms and potential for therapeutically relevant
45 inhibition. *Curr. Enz. Inhib.* 3, 77-94.
46
47 Tricarico, P.M., Kleiner, G., Piscianz, E., Zanin, V., Monasta, L., Crovella, S. & Marcuzzi, A. (2013)
48 Temperature and drug treatments in mevalonate kinase deficiency: an ex vivo study. *Biomed.Res.Int.*
49 2013, 715465.
50
51 Valdar, W.S. (2002) Scoring residue conservation. *Proteins.* 48, 227-241.
52
53 van der Burgh, R., Ter Haar, N.M., Boes, M.L. & Frenkel, J. (2013) Mevalonate kinase deficiency, a
54 metabolic autoinflammatory disease. *Clin.Immunol.* 147, 197-206.
55
56 Van Durme, J., Maurer-Stroh, S., Gallardo, R., Wilkinson, H., Rousseau, F. & Schymkowitz, J. (2009)
57 Accurate prediction of DnaK-peptide binding via homology modelling and experimental data. *PLoS*
58 *Comput.Biol.* 5, e1000475.
59
60

1
2
3 Voynova, N.E., Rios, S.E. & Miziorko, H.M. (2004) Staphylococcus aureus mevalonate kinase: isolation
4 and characterization of an enzyme of the isoprenoid biosynthetic pathway. J.Bacteriol. 186, 61-67.
5

6 Worth, C.L., Preissner, R. & Blundell, T.L. (2011) SDM--a server for predicting effects of mutations on
7 protein stability and malfunction. Nucleic Acids Res. 39, W215-22.
8

9 Worth, C.L., Bickerton, G.R., Schreyer, A., Forman, J.R., Cheng, T.M., Lee, S., Gong, S., Burke, D.F. &
10 Blundell, T.L. (2007) A structural bioinformatics approach to the analysis of nonsynonymous single
11 nucleotide polymorphisms (nsSNPs) and their relation to disease. J.Bioinform Comput.Biol. 5, 1297-
12 1318.
13

14 Yang, D., Shipman, L.W., Roessner, C.A., Scott, A.I. & Sacchettini, J.C. (2002) Structure of the
15 Methanococcus jannaschii mevalonate kinase, a member of the GHMP kinase superfamily.
16 J.Biol.Chem. 277, 9462-9467.
17

18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

Figure legends

Figure1: Structure of human mevalonate kinase showing the location of changes resulting from disease-associated mutations. (a) The dimeric structure of human MVK showing the N- and C-termini and active site cleft of one of the subunits. (b) The sequence of human MVK (GenBank: NP_001107657) showing residues which when altered are associated with mild (green), intermediate (magenta) and severe (red) forms of MVK deficiency. (c) These residues mapped onto the three dimensional structure of a single MVK subunit (left). Each residue associated with disease is shown as a space-filled model. Those associated with the severe form (MA) are coloured red. Magenta indicates residues which are altered in the intermediate form and green residues altered in the mild form (HIDS). The “hot-spot” which includes all the alterations associated with MA is coloured blue in the right hand figure. Parts (a) and (c) were produced using PyMol and (PDB: 2R3V) (Fu et al. 2008).

Figure2: When combined, the intermediate and severe groups of variants show a statistically significant difference from the mild variants in the mean physicochemical score ($p=0.038$).

Supporting Information

Supplementary FigureS1: A sequence alignment generated in Clustal Omega (see Materials and Methods) of 19 verified MVK protein sequences. This alignment was used in analysis of sequence

Supplementary TableS2: The percentage of predictions from 28 prediction tools (see Materials and Methods) for MVK variants predicting a decrease, increase or neutral change in stability. Whether or not the residue occurs in the structural “hotspot” is also tabulated.

45
46 conservation.
47

48 [Supplementary TableS1](#): The complete dataset from the various prediction tools used in this study.
49

45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 [Supplementary TableS3](#): Predicted severities associated with currently uncharacterised variants.

4
5 Variants were predicted to be intermediate/severe (I/S), mild (M) and mild/neutral (M/N) based on
6
7 whether, or not, the altered residue was in the “hotspot” and whether the variant was predicted to
8
9 be substantially less stable than wild-type.
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44

For Peer Review

50
51
52 [Supplementary TableS2](#): The percentage of predictions from 28 prediction tools (see Materials and
53
54 Methods) for MVK variants predicting a decrease, increase or neutral change in stability. Whether
55
56 or not the residue occurs in the structural “hotspot” is also tabulated.
57
58
59
60

1 (a)

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19 (b)

20

21

22

23

24

25

26

27

28

29

30

31 (c)

32

33

34

35

36

37

38

41

42

43

44

45

46

47

48

49

50

51

52

53

54

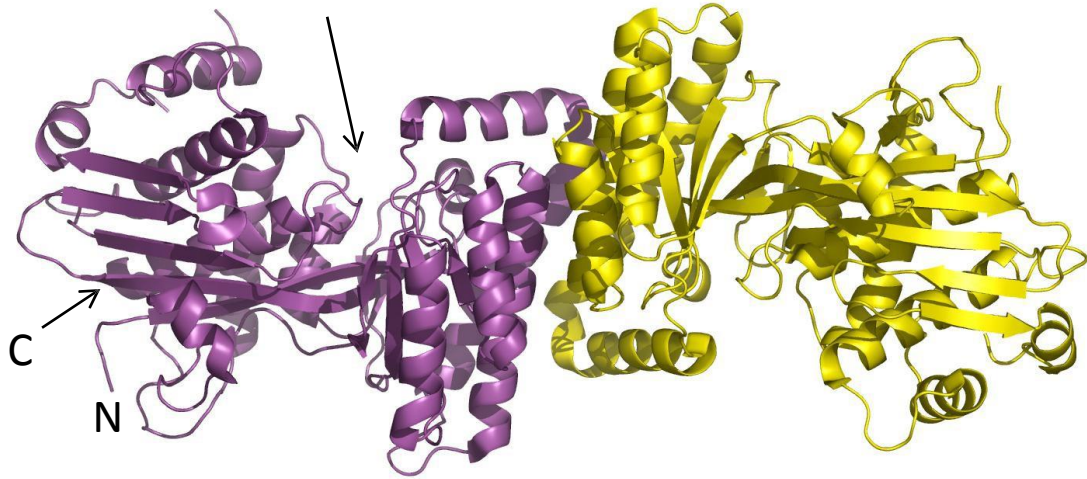
55

56

57

58

Active site cleft



19 (b)

1	MLSEVLLVSA	PGKVILHGEH	AVVHGKVALA	VSLNLR ^L RTF ^L R	LQPHSNGKVD	LSLPNIGIKR
61	AWDVARLQSL	DTSFLEQGDV	TTPTSEQVEK	LKEVAGLPDD	CAVTERLAVL	AFLYLY ^L LS ^I C
121	RKQRALPSLD	IVVWSELPPG	AGLGSSAAYS	VCLAAALLTV	CEEIPN ^L PKD	GDCVNRWTKE
181	DLELINKWAF	QGERMIHGNP	SGVDNAVSTW	GGAL ^R YHQK	ISSLKRSPAL	QILL ^L TNTKVP
241	R ^N TRAL ^L VAG ^V	RNRL ^L KFPEI	VAP ^L LL ^L TS ^I DA	ISLE ^C ERV ^L LG	EMGEAPAPEQ	YLV ^L LEEL ^I IDM
301	NQHHLNAL ^L GV	GHASLDQLCQ	VTRAR ^L GLH ^S K	LTGAGGG ^L CG	ITLLKPGLEQ	PEVEATKQAL
361	TSCGFDCLET	SIGAP ^L GVSIH	SATSLDSRVQ	QALDGL		

31 (c)

32

33

34

35

36

37

38

41

42

43

44

45

46

47

48

49

50

51

52

53

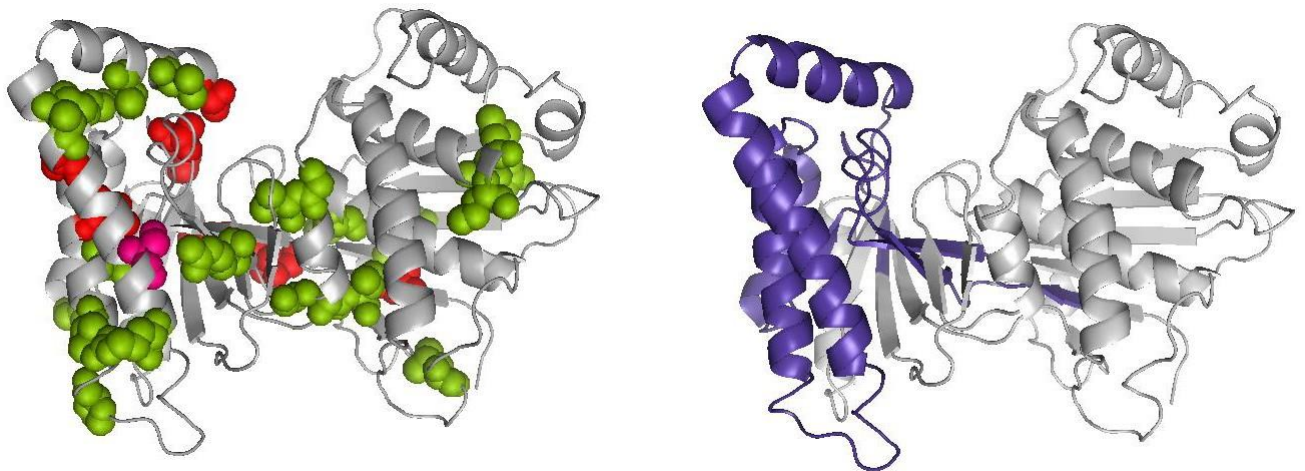
54

55

56

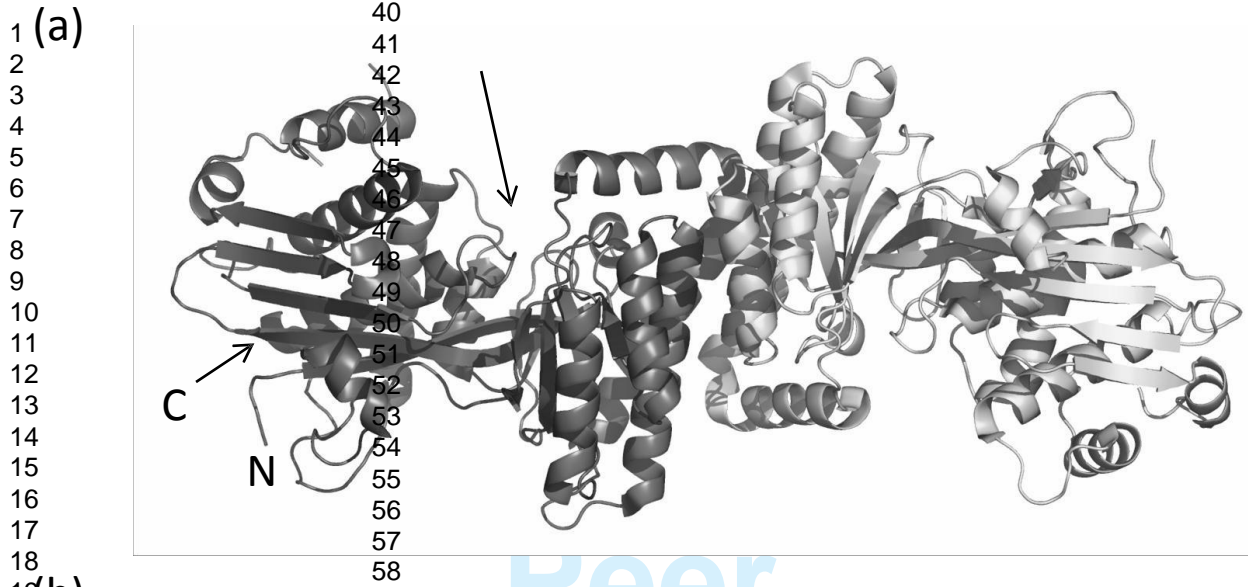
57

58



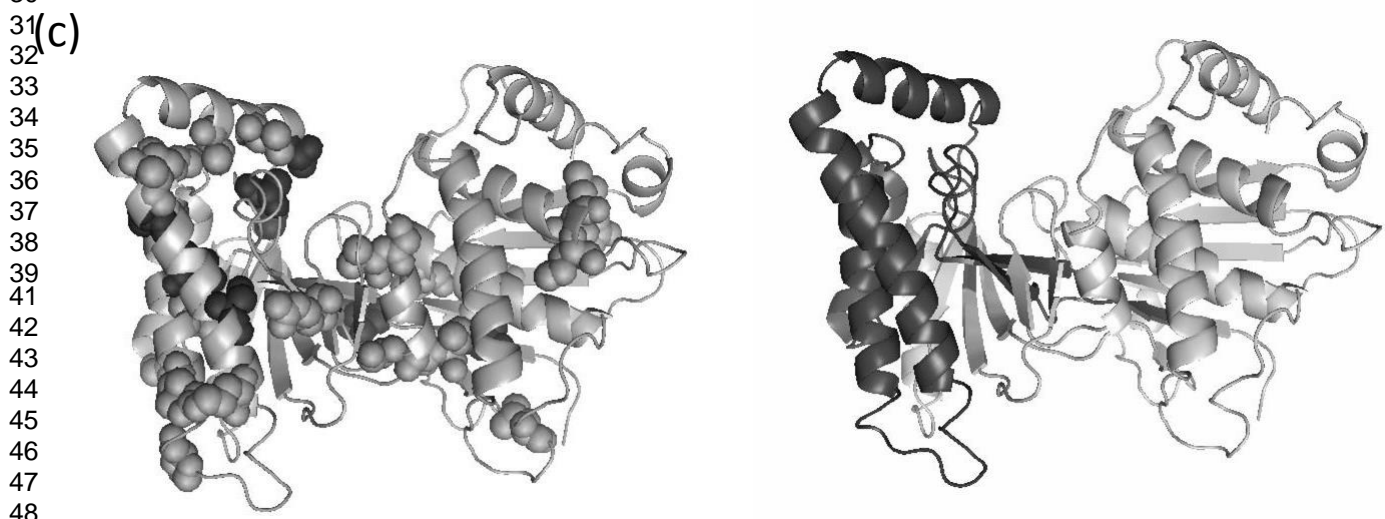
39
40

45
46
47
48
49
50
51
52
53
54
55
56
57
58



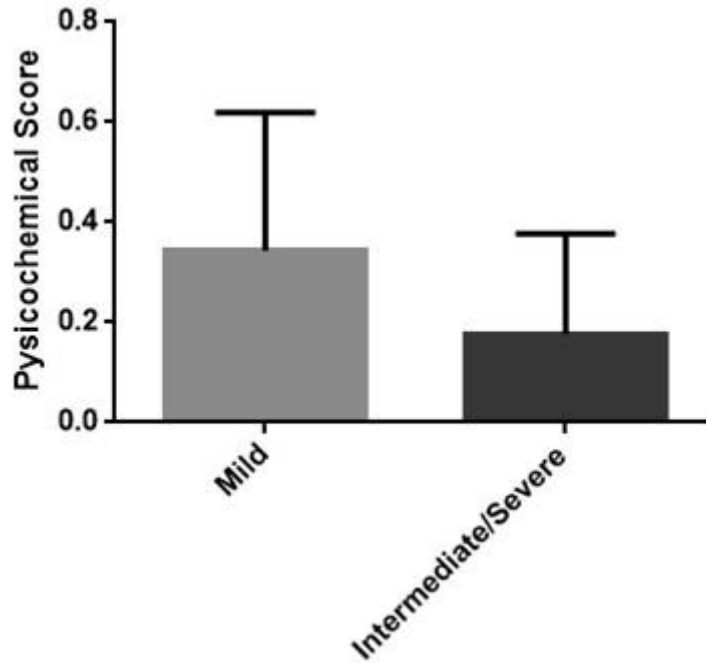
19 (b)

21	1	MLSEVLLVSA	PGKVILHGEH	AVVHGKVALA	VSLNLRTEFLR	LQPHSNGKVD	LSLPNIGIKR
22	61	AWDVARLQSL	DTSFLEQGDV	TTPTSEQVEK	LKEVAGLPDD	CAVTERLAVL	AFLYLYLSIC
23	121	RKQRALPSLD	IVVWSELPPG	AGLGSSAAYS	VCLAAALLTV	CEEIPNPLKD	GDCVNRWTKE
24	181	DLELINKWAF	QGERMIHGPN	SGVDNAVSTW	GGALRYHQGK	ISSLRSPAL	QILLTNTKVP
25	241	RNTRALVAGV	RNRLKFPEI	VAPLLTSIDA	ISLECERVLG	EMGEAPAPEQ	YLVLEELIDM
26	301	NQHHLNALGV	GHASLDQLCQ	VTRARGLHSK	LTGAGGGGCG	ITLLKPGLEQ	PEVEATKQAL
27	361	TSCGFDCLET	SIGAPGVSIH	SATSLDSRVQ	QALDGL		



32
33
34
35
36
37
38
39
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58

Active site
cleft



Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58

Annals of Human Genetics

Table 4. Showing the results from all bioinformatic and structural analysis on known and untagged variants of MX

Table with columns for variant ID, coordinates, gene, and various bioinformatic/structural analysis results. The table is extremely dense and contains a large volume of data points for each variant.



52
53
54
55
56
57
58
59
60

41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58