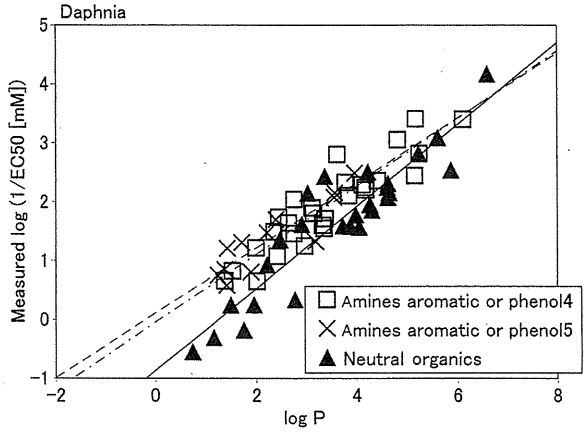

Figure 1. The correlation between log P and the measured toxicity values of chemicals used in KATE as a daphnia end-point. The dotted-dashed, dashed and bold lines are the QSAR equations of *amines aromatic or phenols4*, *amines aromatic or phenols5*, and *neutral organics*, respectively.

number of chemical substances in the *unacceptable* group. For example, the fish *hydrocarbons aromatic* class had 43 reference data, $r^2 = 0.826$, RMSE = 0.368, and only one unacceptable chemical. In other words, 98% of the chemicals were classed as acceptable. On the other hand, the fish *dinitrobenzene* class contained 12 reference data, $r^2 = 0.331$, RMSE = 0.669, and three unacceptable chemicals. In this case, 75% of the chemicals were thus acceptable.

As shown in Tables 1 and 2, each of the classes with $r^2 \geq 0.7$, RMSE ≤ 0.5, and $n > 5$, e.g., the fish *hydrocarbon aromatic* class, had a sufficiently high q^2. Such classes showed QSAR equations similar to those of *neutral organics*. Thus the toxicity of such classes could be explained mainly by the narcotic effect of the chemicals. However, the daphnia *amines aromatic or phenols4* and *amines aromatic or phenols5* groups had a larger intercept b in the QSAR equations than *neutral organics* with a small log P value (see Figure 1). These classes can be explained in terms of polar narcosis or narcosis II [17]. Narcosis II is known to be more toxic than baseline toxicity, i.e., than *neutral organics*, non-polar narcosis, narcosis I, or less inert, as explained by Verhaar et al. [18].

In some cases the q^2 values were much smaller than those of r^2. QSAR equations based on fewer than six reference data require a greater number of reference chemicals.

4.2 *External validation*

Tables 3 and 4 list the statistical data of the TIMES, ECOSAR, and KATE with or without the applicability domains. The complete results are given in Appendix 5 of the supplementary material. First, we will focus on the TIMES, ECOSAR, and all the KATE results, without considering any applicability domains. In fish, the determination coefficient, r^2, and RMSE using KATE ($r^2 = 0.868$ and RMSE = 0.658) were larger and smaller, respectively, than those using TIMES ($r^2 = 0.751$ and RMSE = 0.935) and than by ECOSAR ($r^2 = 0.790$ and RMSE = 0.869). For daphnia, RMSE using KATE (0.993) was smaller than that using TIMES (1.404) and ECOSAR (1.364). However, r^2 using KATE (0.662) showed no noticeable advantage over that by TIMES (0.668) or ECOSAR (0.699).

Table 3. Statistical information comparing measured and calculated fish $\log(1/LC_{50}[\text{mM}])$ of 287 test set compounds. The complete results are shown in Appendix 5-1.

			KATE[*2]					
	TIMES[*1]	*ECOSAR*[*2]	*All*[*3]	*log P*[*4]	*C(1)*[*5]	*C(2)*[*6]	*log P*[*4] *C(1)*[*5]	*log P*[*4] *C(2)*[*6]
Chemicals[*7]	274	242	274	207	152	192	111	144
Predicted[*8]	274	259	318	252	187	233	145	179
r^2	0.751	0.790	0.868	0.833	0.901	0.890	0.886	0.866
RMSE	0.935	0.869	0.685	0.641	0.644	0.655	0.588	0.617
Under[*9] [%]	11.3	10.0	4.7	5.2	5.3	5.6	2.8	3.9
Over[*10] [%]	5.1	8.1	7.2	6.7	8.0	6.9	8.3	7.3

Notes:
[*1]Each chemical is identified by one QSAR class.
[*2]When a chemical is found to belong to more than one QSAR class, all the estimated data are adopted. If only the name of the class is available, such data are omitted.
[*3]Both in-domain and out-of-domain data for log *P* and *C*-judgements are included.
[*4]In-domain of log *P*-judgement.
[*5]In-domain of *C*-judgement is defined as all substructures of a test chemical being found in reference chemicals in the class.
[*6]In-domain of *C*-judgement defined as all substructures of a test chemical being in reference chemicals in either *Neutral organics* or the class.
[*7]The number of compounds that can be predicted.
[*8]The total number of the predicted values by using the training sets. Some chemicals belong to more than one class, and thus *Predicted* is larger than *Chemicals*. r^2, RMSE, Under and Over were calculated based on the *Predicted* number.
[*9]Fractions (%) of the underestimated chemicals. Underestimation is defined as [calculated $\log(1/LC_{50})$ – measured $\log(1/LC_{50})$] <-1.
[*10]Fractions (%) of the overestimated chemicals. Overestimation is defined as [calculated $\log(1/LC_{50})$ – measured $\log(1/LC_{50})$] >1.

Table 4. Statistical information between measured and calculated *Daphnia* $\log(1/EC_{50}[\text{mM}])$ for 98 test set compounds. The complete results are shown in Appendix 5-2.

			KATE[*2]					
	TIMES[*1]	*ECOSAR*[*2]	*all*[*3]	*log P*[*4]	*C(1)*[*5]	*C(2)*[*6]	*log P*[*4] *C(1)*[*5]	*log P*[*4] *C(2)*[*6]
Chemicals[*7]	93	82	94	58	43	55	25	33
Predicted[*8]	93	85	102	66	46	61	31	39
r^2	0.668	0.699	0.662	0.732	0.793	0.686	0.807	0.801
RMSE	1.404	1.364	0.993	0.784	0.799	0.968	0.639	0.689
Under[*9] [%]	21.5	14.1	9.8	1.5	6.5	8.2	0.0	0.0
Over[*10] [%]	11.8	18.8	14.7	15.2	6.5	11.5	6.5	10.3

Notes: As in Table 3.

Since reference data for the daphnia end-point (258 chemicals) numbered only half of those for fish (535 chemicals), the reference data for each QSAR equation for daphnia would therefore be less satisfactory for predicting toxicity. The addition of reference data and a change in the classification rules can recover the values of the statistical data.

A fraction of $\log(1/LC_{50})$ with an underestimation of less than -1 indicated that, compared with KATE, TIMES and ECOSAR tended to underestimate the toxicities of both fish and daphnia. On the other hand, a fraction of $\log(1/LC_{50})$ showing an overestimation of more than 1 indicated that, compared with TIMES, ECOSAR and KATE tended to overestimate toxicity in both fish and daphnia. Considering these under- and over-estimation fractions, we find that KATE gives a higher predictive ability in acute *Oryzias latipes* and *Daphnia magna* toxicity tests than does TIMES or ECOSAR. If the alert: *Out of domain*, in TIMES, and the applicable $\log P$ range in ECOSAR are considered rigidly, the correlation between measured and calculated toxicity is improved in TIMES and ECOSAR.

Secondly, in fish, the RMSE of one of any in-domains was smaller than if domains were not considered. However, the r^2 in-domain of $\log P$ showed no particular improvement. For daphnia, r^2 and RMSE for one of any in-domains were larger and smaller, respectively, than those without considering domains. In the present study, either the descriptor and/or structural domains were related to the reduction of RMSE and the fraction of underestimated chemicals, especially if both domains were considered simultaneously. Additionally, the stricter structural domain C(1) (shown in Tables 3 and 4) demonstrated better predictive performance than the structural domain C(2). The systematic study of the domain based on the atom-centred fragment (ACF) approach by Kuhne et al. [14] showed that the ACF varied with respect to its size in terms of the path length, and the ACF match mode was specified in terms of degree of strictness. They also demonstrated a clear relationship between predictive performance and the levels of the ACF definition and match mode [14]. Even though the definition of substructures for the domain are different, the improvement by using *C*-judgement is similar in concept to that using the ACF approach. Thus, the $\log P$ range of the equation and *C*-judgement are useful for assessing the applicability of the QSAR results.

5. Summary

We have reported on the KATE system, encompassing a full list of classifications of the QSAR equations and KATE validations. In the KATE system chemicals are classified by their substructure. The QSAR equations express the correlation between $\log P$ and $\log(1/LC_{50})$ or $\log(1/EC_{50})$ of a chemical by simple linear regression analyses. The classes of QSAR equations are characterised by fragments of chemicals, except for the *neutral organics* class. The descriptor and structure domains, $\log P$ and *C*-judgements, in KATE were also introduced.

The cross-validation of the KATE system showed that QSAR equations with higher r^2 and lower RMSE with $n > 5$ gave a reliably higher q^2 than the other QSAR equations in KATE, meaning they had better predictive ability. A comparison of KATE, TIMES, and ECOSAR revealed that KATE was more accurate, due to end-point dependence. The use of $\log P$ and the *C*-judgement improved the statistical data. Thus the KATE system is a powerful tool for predicting acute toxicity in *Oryzias latipes* and *Daphnia magna* when the $\log P$ and *C*-judgement can be confirmed. Also, KATE has the potential to be useful in risk assessment.

The next topics in QSAR development will be to consider the reactivity of chemicals, and to include multi-regression analysis. The quantum chemical parameters, such as partial charges, are candidates for additional descriptors. Other ways of significantly

increasing the reliability of toxicity prediction will be to improve the classification of the substructures, increase the reference data in a QSAR equation, and to refine the *C*-judgement.

Acknowledgements

KATE was researched and developed by the Research Center for Environmental Risk at the NIES, under contract to the Japanese MoE between 2004 and 2008. We also wish to thank the US EPA for permission to use KOWWIN in KATE on PAS, the standalone version of the KATE system. We are grateful to Mr K. Hasunuma and Ms K. Sugiyama for their support and encouragement with the KATE publication.

References

[1] T.I. Netzeva, M. Pavan, and A.P. Worth, *Review of (quantitative) structure–activity relationships for acute aquatic toxicity*, QSAR Comb. Sci. 27 (2008), pp. 77–90.

[2] OECD, *Report on the regulatory uses and applications in OECD member countries of (quantitative) structure–activity relationship [(Q)SAR] models in the assessment of new and existing chemicals*, Environment Health and Safety Publications Series on Testing and Assessment, No. 58, OECD, Paris, 2006.

[3] MoE, Japan ecotoxicity tests data. Available at http://www.env.go.jp/chemi/sesaku/02e.pdf

[4] KATE. Available at http://kate.nies.go.jp Copyright (C) 2008–2009 Ministry of the Environment, Government of Japan, all rights reserved. It is cautioned that these QSAR results may not be used as ecotoxicity test results required for MoE submissions in compliance with the CSCL.

[5] O.G. Mekenyan, S.D. Dimitrov, T.S. Pavlov, and G.D. Veith, *A systematic approach to simulating metabolism in computational toxicology. I. The TIMES heuristic modelling framework*, Curr. Pharm. Des. 10 (2004), pp. 1273–1293.

[6] S.D. Dimitrov, O.G. Mekenyan, G.D. Sinks, and T.W. Schultz, *Global modeling of narcotic chemicals: Ciliate and fish toxicity*, THEOCHEM. 622 (2003), pp. 63–70.

[7] U.S. EPA, ECOSAR™. Available at http://www.epa.gov/oppt/newchems/tools/21ecosar.htm. See also http://www.epa.gov/oppt/newchems/tools/ecosartechfinal.pdf

[8] US EPA, fathead minnow database. Available at http://www.epa.gov/med/Prods_Pubs/fathead_minnow.htm

[9] C.L. Russom, S.P. Bradbury, S.J. Broderius, D.E. Hammermeister, and R.A. Drummond, *Predicting modes of toxic action from chemical structure: Acute toxicity in the fathead minnow (Pimephales promelas)*, Environ. Toxicol. Chem. 16 (1997), pp. 948–967.

[10] OECD, *Report of the OECD Workshop on quantitative structure activity relationships (QSARs) in aquatic effects assessment, Environment Monographs, No. 58*, OECD, Paris, 1992.

[11] Clog P, Daylight Chemical Information Systems, Inc. Available at http://www.daylight.com/dayhtml/doc/clogP/index.html. The underlying program, CLOG P, is copyrighted by Pomona College and BioByte, Inc., of Claremont, CA.

[12] S. Dimitrov, G. Dimitrova, T. Pavlov, N. Dimitrova, G. Patlewicz, J. Niemela, and O. Mekenyan, *A stepwise approach for defining the applicability domain of SAR and QSAR models*, J. Chem. Inf. Model. 45 (2005), pp. 839–849.

[13] T.W. Schultz, M. Hewitt, T.I. Netzeva, and M.T.D. Cronin, *Assessing applicability domains of toxicological QSARs: Definition, confidence in predicted values, and the role of mechanisms of action*, QSAR Comb. Sci. 26 (2007), pp. 238–254.

[14] R. Kuhne, R.U. Ebert, and G. Schuurmann, *Chemical domain of QSAR models from atom-centered fragments*, J. Chem. Inf. Model. 49 (2009), pp. 2660–2669.

[15] US EPA, KOWWIN™. Available at http://www.epa.gov/opptintr/exposure/pubs/episuite.htm
[16] E.M. Hulzebos and R. Posthumus, *(Q)SARs: Gatekeepers against risk on chemicals?*, SAR QSAR Environ. Res. 14 (2003), pp. 285–316.
[17] G.D. Veith and S.J. Broderius, *Rules for distinguishing toxicants that cause Type-I and Type-II Narcosis syndromes*, Environ. Health Perspect. 87 (1990), pp. 207–211.
[18] H.J.M. Verhaar, C.J. Vanleeuwen, and J.L.M. Hermens, *Classifying environmental pollutants: 1. Structure–activity relationships for prediction of aquatic toxicity*, Chemosphere 25 (1992), pp. 471–491.

Science of the Total Environment 409 (2011) 3073–3077

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Short Communication

Estimation of daily inhalation rate in preschool children using a tri-axial accelerometer: A pilot study☆

Junko Kawahara [a,*], Shigeho Tanaka [b], Chiaki Tanaka [c], Yasunobu Aoki [a], Junzo Yonemoto [a]

[a] Research Center for Environmental Risks, National Institute for Environmental Studies, Tsukuba, Japan
[b] Health Promotion and Exercise Program, National Institute of Health and Nutrition, Tokyo, Japan
[c] Division of Integrated Sciences, J.F. Oberlin University, Tokyo, Japan

ARTICLE INFO

Article history:
Received 26 October 2010
Received in revised form 27 March 2011
Accepted 1 April 2011

Keywords:
Daily inhalation rate
Exposure assessment
Preschool children
Accelerometer
Activity monitoring
Time-activity

ABSTRACT

The activity of 5- to 6-year-old Japanese children (n=29) was monitored for 3 consecutive days, including one weekend day, using an ActivTracer tri-axial accelerometer. The daily inhalation rate and time spent in sedentary, light, or moderate to vigorous levels of physical activity (MVPA) were estimated from the accelerometer measurements based on previously developed regression equations. The 3-day mean daily inhalation rate (STPD) was estimated at $8.3 \pm 1.4\ m^3\ day^{-1}$ in 10 subjects who completed 3 days of monitoring. The time spent in sedentary, light, or MVPA each day was 320, 415, and 81 min day^{-1}, respectively. Analysis of between-day reliability indicated that 3 days of monitoring with the ActivTracer tri-axial accelerometer provided an acceptable estimate of daily inhalation rate (intra-class correlation coefficient [ICC]=0.892), but low to moderate reliability for the time spent in different levels of activities (ICC=0.43 to 0.58). We observed a significant difference in the daily inhalation rate between weekdays and the weekend day, possibly due to differences in time spent in MVPA. This finding suggests that a weekend day should be included to obtain more reliable estimates of daily inhalation rate using an accelerometer.

1. Introduction

The inhalation rate is an essential factor for estimating the inhaled doses of air pollutants. In recent years, consideration of particular vulnerability and patterns of exposure in children has been an issue in the assessment of health risks associated with hazardous environmental pollutants (WHO, 2006). The average daily inhalation rate of 8.7 $m^3\ day^{-1}$ established for U.S. children aged between 1 and 12 years (U.S. EPA, 1997) is currently proposed for health risk assessment in Japanese children to calculate the inhaled doses of pollutants (MHLW, 2007). However, differences in body size and daily food intake between American and Japanese children would be expected to introduce bias in estimations of daily inhalation rate in Japanese children. An approach based on energy intake rate has also been reported to overestimate the daily inhalation rate in younger age groups by 7% to 35% compared with those evaluated by the doubly-labeled water (DLW) method (Brochu et al., 2006b).

The metabolic energy expenditure approach and the time–activity–ventilation (TAV) approach are the methods used to estimate daily inhalation rate in humans (Arcus-Arth and Blaisdell, 2007; Stifelman, 2007; Brochu et al., 2006a; U.S. EPA, 2006; Allan and Richardson, 1998; Layton, 1993; ICRP, 1975). Of the energy expenditure approaches available, a physiological approach that uses daily energy expenditure measured by the 'gold standard' DLW technique (IOM, 2005) provides the most accurate estimate of the daily inhalation rate in individuals during daily life. However, the ^{18}O isotope used in this technique is very expensive and collection of urine samples over a 2-week period is usually required, resulting in the test being unsuitable for studies with a large number of subjects. On the other hand, the TAV method is a traditional approach that estimates the daily inhalation rate using existing data of minute respiratory ventilation rate ($\dot{V}_E$) of various physical activities performed during daily life. While the advantage of this approach is that the $\dot{V}_E$ data used in the calculations are actual measurements, a critical issue is the limited availability of data on $\dot{V}_E$ values for a variety of children's activities. In order to fill these gaps in the data, Layton (1993) developed a time–activity energy expenditure approach that calculated daily inhalation rate using metabolically derived $\dot{V}_E$ as a product of basal metabolic rate (BMR), metabolic equivalent (MET) for an activity of interest, ventilatory equivalent (VQ),

☆ Sources of support: This study was supported in part by the Ministry of Environment of Japan. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the ministry.

* Corresponding author at: 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan. Tel.: +81 29 850 2695; fax: +81 29 850 2920.
E-mail address: jkawa@nies.go.jp (J. Kawahara).

0048-9697/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.scitotenv.2011.04.006

and oxygen uptake rate. The strength of this "MET approach" is that it uses a $\dot{V}_E$ that corresponds to a variety of physical activities that people engage in during daily life, which can be estimated from MET values abundantly available in the literature. However, the limitation of this approach is that the MET value should not be used to calculate the cost of the children's activities (Torun, 1983; Puyau et al., 2004). Self-report based time–activity data used in the TAV and MET approaches (Allan and Richardson, 1998; U.S. EPA, 2006) also have limitations in objectivity and accuracy, especially in young children (Bender et al., 2005; Lichtman et al., 1992).

To overcome the issues described above, we introduced a new approach in a previous study that used an accelerometer to estimate $\dot{V}_E$ that corresponded to a range of physical activity levels in children during daily life (Kawahara et al., 2011). In the present study, we conducted 3 days of continuous monitoring of physical activity in preschool children during daily life. Using the regression equations developed in our previous study we estimated daily inhalation rate and time spent at different levels of physical activity intensity from minute-by-minute ActivTracer accelerometer measurements. The objective of the present study was to examine the between-day reliability of these summary variables over 3 days, including one weekend day. We also assessed the magnitude of under- or over-estimation of our daily inhalation estimates by comparing the values with published daily inhalation rates derived from daily energy expenditure measured by the DLW method.

2. Methods

2.1. Subjects

The participants in the present study were 29 Japanese preschool children aged 5 to 6 years (mean ± SD, 6.2 ± 0.2 years; range, 5.9–6.8 years). The boys (n = 16) and girls (n = 13) were recruited from a single kindergarten class in a suburb of Tokyo. Of those participants, 26 children had participated in our previous laboratory exercise test, in which we calibrated measurements with the ActivTracer tri-axial accelerometer against $\dot{V}_E$ measured with the Douglas bag method (Kawahara et al., 2011). The mean height of the participants was 116 ± 5 cm (range, 103–125 cm) and mean body weight was 20.0 ± 2.3 kg (range, 16.3–25.5 kg). No significant gender differences in body weight or height were observed (*t*-test, $P<0.05$). The experimental procedures and purpose and protection of personal information were explained thoroughly to the parents of each participant, and written consent was obtained before monitoring was initiated. The experimental protocol was approved by the ethics committee of the National Institute for Environmental Studies.

2.2. Data collection

The physical activity of the children was monitored for 3 consecutive days, including one weekend day (i.e., from Thursday morning to Sunday morning) during the period December 2006 to March 2007. The monitoring period was determined based on the results of previous studies (Trost et al., 2000; Janz et al., 1995) that indicated a 3-day period was the minimum duration required for reliable estimates of typical physical activity in children using an accelerometer. In addition, Trost et al. (2000) suggested including weekend days to ensure reliable assessment of physical activity in children using an accelerometer. In the present study, Thursday (day 1) and Friday (day 2) served as the weekdays, and Saturday (day 3) served as the weekend day. On weekdays during the monitoring period, the children attended kindergarten as usual. Their activity was monitored with an ActivTracer tri-axial accelerometer (AC-210A, 50 mm × 70 mm × 20 mm, 60 g, GMS Inc., Tokyo, Japan). The ActivTracer detects movements in the anteroposterior (*x*-axis), mediolateral (*y*-axis), and vertical (*z*-axis) directions. The output measure of the ActivTracer is the average of absolute values for acceleration in each direction and the synthetic values of the 3 axes (vector magnitude) for a time interval defined by the user. The monitor was set to record body acceleration at 60-second intervals, and was contained in a small nylon pouch worn on the subject's hip attached to an adjustable belt. The subjects were allowed to detach the monitor when sleeping, bathing, showering, and swimming. The parents were instructed to help their child wear the monitor continuously during all waking hours for 3 consecutive days. They were also given a time–activity log sheet to record the time the subjects detached the monitor and the activities the subjects engaged in while monitoring was discontinued. During the time the subjects were at the kindergarten, the time–activity log was completed by the study staff.

2.3. Data reduction

Periods of 5 min or longer of continuous non-detection of body acceleration and no record of the monitor being detached were considered as non-wearing times and were not included in the calculation of total minutes of monitoring. Subjects who had non-wearing times longer than 60 min per day, with the exception of times for sleeping, bathing, showering, or swimming, were regarded as poorly compliant with the monitoring protocol. Only subjects with 3 complete days of monitoring data were included in the estimation of daily inhalation rate and time spent at different levels of physical activity.

2.4. Estimation of the daily inhalation rate

The daily inhalation rate (Standard temperature, Standard Pressure, and Dry, STPD) was calculated by summing the $\dot{V}_{E,\,STPD}$ for each minute estimated from the ActivTracer measurements during the 24 hour period. $\dot{V}_{E,\,STPD}$ for each minute was estimated from the synthetic acceleration (AC_{xyz}) using the following Eqs. (1) and (2) developed in our previous study based on data obtained from 5 to 6 year-old children including subjects in the current study

$$\dot{V}_{E,\,STPD\,i} = 0.00086 \times AC_{xyz} + 0.20 \tag{1}$$

$$\dot{V}_{E,\,STPD\,i\ other\ than\ walking} = 0.00094\ AC_{xyz} + 0.35 \tag{2}$$

One of the two equations specific for the type of physical activity was used. This is because there are different accelerometer measurements – $\dot{V}_E$ relationships by the type of activity. The details of these procedures have been reported elsewhere (Kawahara et al., 2011). Eq. (1) was used as the default equation to estimate $\dot{V}_{E,STPD}$. If the synthetic acceleration of the activities of interest was within the range of 96–754 mG min^{-1}, the activities were classified as either 'walking' or 'other than walking' types using the discriminating Eq. (3) described below:

$$F = -3.31 + 5.98 \times \left(AC_z / AC_{xy}\right) - 0.017 \times AC_z \tag{3}$$

In this equation, F is a discriminant function, AC_z is vertical acceleration, and AC_{xy} is horizontal acceleration measured with the ActivTracer. If F was <0.207, the activity was classified as a type 'other than walking'. For such an activity, $\dot{V}_{E,STPD}$ was estimated from the synthetic acceleration using Eq. (2). For activity classified as 'walking,' $\dot{V}_{E,\,STPD}$ was estimated using Eq. (1).

Based on the previous study (Kawahara et al., 2011), the values of $\dot{V}_{E,\,STPD}$ during periods when the subjects had detached the monitor were assumed to be as follows: During sleeping, 0.16 L kg^{-1} min^{-1}, derived by dividing 0.18 L min^{-1} at rest and lying quietly by 1.1, on the assumption that the metabolic rate during sleep is 10% lower than at rest (IOM, 2005), during use of the toilet and watching TV,

0.21 L kg^{-1} min^{-1}; during eating, 0.25 L kg^{-1} min^{-1}; during dressing/undressing and bathing/showering, 0.51 L kg^{-1} min^{-1}; and playing board games, 0.26 L kg^{-1} min^{-1}(Kawahara et al., 2011). As reliable data regarding $\dot{V}_E$ values or the physical activity ratio (PAR) for swimming in young children are not currently available, data for subjects who swam during the monitoring period were excluded from the analysis in the present study.

2.5. Estimation of time–physical activity distribution

The daily duration spent either sedentary, in light level of physical activity (LPA), or in moderate to vigorous level of physical activity (MVPA) was estimated based on minute-by-minute accelerometer measurements and partly on the time–activity records. As a measure of the intensity of physical activity, we used PAR, the energy cost of an activity per unit time, which is defined as a multiple of BMR per minute (FAO, 2004). According to the classification of Puyau et al., 2004, PAR <1.5 = sedentary, PAR ≥ 1.5 and <3.0 = LPA, and PAR ≥ 3.0 = MVPA. The cut-off levels between sedentary and LPA and LPA and MVPA were 71 mG min^{-1} and 412 mG min^{-1}, respectively, and for types of activity 'other than walking,' the cut point between LPA to MVPA was 218 mG min^{-1} (Kawahara et al., 2011). Physical activity including using the toilet, watching TV, eating, and playing board games was considered sedentary (Tanaka et al., 2007; Kawahara et al., 2011), while dressing/undressing and bathing/showering were classified as LPA (Taylor et al., 1948; Yamamura et al., 2003).

2.6. Statistical analysis

Intra-class correlation coefficients (ICC) were calculated using repeated measures analysis of variance (ANOVA) in order to evaluate the between-day reliability of daily inhalation rate, and the average time spent over the 3 days either sedentary, in LPA, or in MVPA. Repeated measures ANOVA was carried out to detect significant group mean differences in average daily inhalation rate and time spent in each level of physical activity among the 3 days of monitoring with day of monitoring as the within-subject variable. If the result of the analysis was significant, Tukey's test was used to determine which days were different. Bland–Altman plots (Bland and Altman, 1986) were prepared to assess the agreement in daily inhalation rate between weekdays and the weekend day. Statistical significance was defined as $P<0.05$. All statistical analyses were performed with SPSS (ver. 15.0 for Windows, SPSS Inc. Tokyo, Japan).

3. Results

Of the total of 87 days of monitoring, complete 1-day data were available for 60 days. Complete 3-day monitoring data sets were obtained from 15 subjects. Of these, 5 subjects reported they had carried out swimming activities for 60 ± 29 min day^{-1} during the monitoring period. The 3-day mean monitoring time for the 10 subjects with complete 3-day monitoring data was 779 ± 51 min day^{-1}, and the daily inhalation rate (STPD) was estimated at 8.3 ± 1.4 m^3 day^{-1}. The 3-day mean daily times spent sedentary, in LPA, or MVPA were 320 ± 51 min day^{-1}, 415 ± 47 min day^{-1}, and 81 ± 34 min day^{-1}, respectively (Table 1). The ICC for the average daily inhalation rate (STPD) for 3 days was 0.961 ($P<0.001$). The single-day reliability coefficient for the daily inhalation rate (STPD) was 0.892 ($P<0.001$). The average daily inhalation rates (STPD) were significantly different between the 3 days ($P=0.01$). Fig. 1 shows the Bland–Altman plots of the differences for each subject against the mean daily inhalation rate (STPD). The mean difference in daily inhalation rate (STPD) between days 1 and 2 was 0.1 m^3 day^{-1} (95% CI: −1.03 and 1.23 m^3 day^{-1}). The mean difference in the daily inhalation rate (STPD) between days 1 and 3 was 0.63 m^3 day^{-1} (95% CI: −0.56 and 1.73 m^3 day^{-1}) and 0.54 m^3 day^{-1} (95% CI: −0.58 and 1.65 m^3 day^{-1}) between days 2 and 3. The between-day reliability for mean daily time spent in LPA was acceptable, with an ICC = 0.63 ($P=0.03$), while it was

Table 1
Mean (SD) for acceleration measurements, daily inhalation rates (STPD), and time spent in various activities estimated from accelerometer measurements in subjects with 3 complete days of monitoring data (N = 10).

Variable	Day 1	Day 2	Day 3	3 days mean
Average acceleration (mG min^{-1})[a]	155 ± 53	161 ± 59	111 ± 40	137 ± 46
Daily inhalation rate (m^3 day^{-1})	8.7 ± 1.6	8.6 ± 1.8	7.8 ± 0.8*	8.3 ± 1.4
Daily time for sleeping (min day^{-1})	609 ± 27	614 ± 40	648 ± 50**	624 ± 33
Daily time for sedentary (min day^{-1})	309 ± 61	311 ± 72	341 ± 77***	320 ± 51
Daily time in a light LPA (min day^{-1})	430 ± 79	418 ± 51	399 ± 57	415 ± 47
Daily time in MVPA (min day^{-1})	92 ± 47	98 ± 52	53 ± 38	81 ± 34

[a] Average acceleration = Σ acceleration measurement/number of monitored minutes.
* Significantly different from day 1 ($P=0.02$) and day 2 ($P=0.03$) based on ANOVA followed by Tukey test.
** Significantly different from day 1 ($P=0.02$) and day 2 ($P=0.04$).
*** Significantly different from day 2 ($P=0.04$).

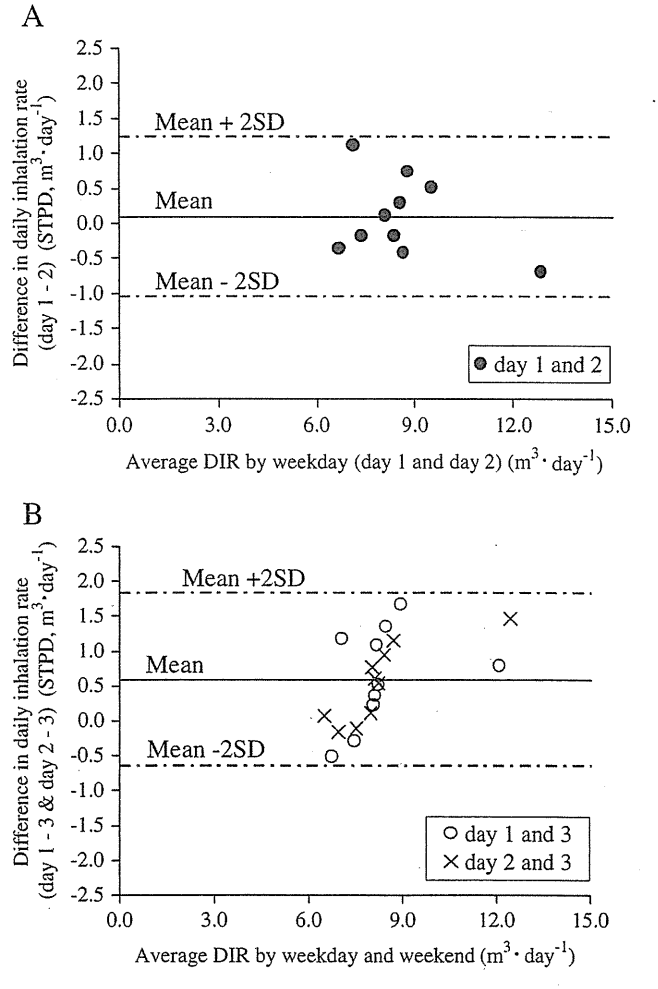

Fig. 1. Bland–Altman plots for the difference in daily inhalation rate (STPD) against mean daily inhalation rate (STPD) by weekday and weekend day. A: difference between days 1 and 2 against mean by days 1 and 2. B: difference between days 1 and 3 and days 2 and 3 against mean by days 1 and 2 and days 2 and 3.

low for sedentary, (ICC = 0.53, $P = 0.07$) and for MVPA (ICC = 0.42, $P = 0.149$).

4. Discussion

Our results indicated that between-day reliability was high for average daily inhalation rate estimated from ActivTracer measurements collected over 3 days, including a weekend day. However, in order to obtain a reliable estimate of daily time spent at different levels of physical activity, more than 3 days is needed. The between-day reliability of daily time spent at different levels of physical activity in this study was low compared with the study of Janz et al. (1995). In that study, the ICC for the percentage of the day spent in sedentary, moderate, or vigorous levels of activity monitored over 3 days in 7- to 15-year-old children using a CSA uni-axial accelerometer were 0.73, 0.70, and 0.71, respectively. This variability between the studies may be due to differences in evaluation of physical activity of children using tri-axial or uni-axial accelerometers (Kawahara et al., 2011) and also differences in the study subjects.

Fig. 2 shows comparison of mean daily inhalation rate (STPD) and distribution of the volume for the 4 categories of physical activity intensity in our subjects by day. Daily inhalation rate in the weekend day was on average 7% lower than during weekdays. The observed difference in the daily inhalation rate between weekdays and weekend days may be due to transition of time spent in MVPA to a lower level of physical activity on weekend days. We consider this result indicates that activity data also needs to be collected during weekend days in order to obtain a more reliable estimation of daily inhalation rate.

Table 2 shows comparison of mean daily inhalation rate expressed at body temperature, ambient pressure, and saturation with water vapor (BTPS) derived from daily energy expenditure measured with the DLW method in previous studies and current study. For comparison with estimates reported in earlier studies, our estimate of the daily inhalation rate of 8.3 $m^3\,day^{-1}$ expressed at STPD, corresponds to 10.1 $m^3\,day^{-1}$ (BTPS). Based on published data on daily energy expenditure measured with the DLW technique, Brochu et al. (2006a) estimated the daily inhalation rate (BTPS) of 5- to 6-year-old boys and girls as 8.6 $m^3\,day^{-1}$ (0.42 $m^3\,kg\,day^{-1}$) and 8.2 $m^3\,day^{-1}$ (0.40 $m^3\,kg\,day^{-1}$), respectively. Stifelman (2007), using the same approach as Brochu et al., reported the daily inhalation rate (BTPS) for 6-year-old boys and girls in an active day (physical activity level = 1.6 to 2.5 (IOM, 2005)), as 9.8 and 9.3 $m^3\,day^{-1}$, respectively (0.57 $m^3\,kg\,day^{-1}$ for both genders; body weight of 22.1 kg for boys and 21.3 kg for girls (U.S. EPA, 2008)). It should be noted that these studies used a VQ value of 27 derived from adult data (Layton, 1993) for estimating the daily inhalation rate of children. The VQ value is evidently lower than the observed value of 30 to 40 during sedentary to vigorous levels of physical activity in 6-year-old children (Kawahara et al., 2010). If we consider that the use of an adult VQ value would lead to a 30% under-estimation of the daily inhalation rate in young children, our estimation is therefore comparable to the rate measured in 5- to 6-year-old children by Brochu et al.

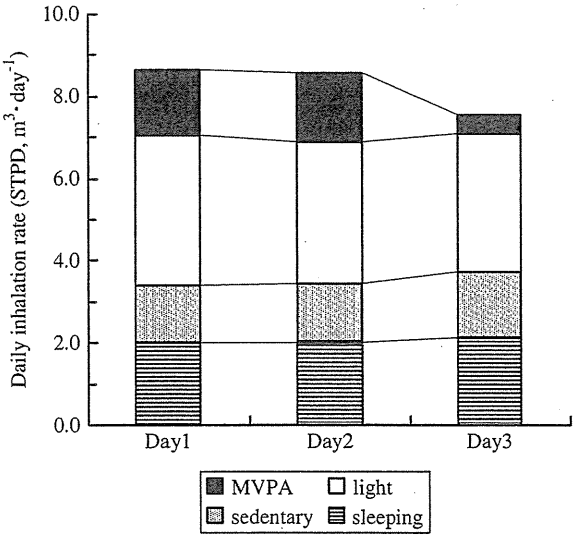

Fig. 2. Daily inhalation rate (STPD) and distribution of the volume for the 4 categories of physical activity intensity: sleeping, sedentary, light level of physical activity (LPA) and moderate to vigorous level of physical activity (MVPA).

Table 2
Comparison of the daily inhalation rate (BTPS) estimated in the current study with published values derived from daily energy expenditure measured with the DLW method.

Study	Age (years)	Gender	Daily inhalation rate (BTPS)[a]	
			(m^3 day)	(m^3 kg day^{-1})
Brochu et al. (2006a,b)[b]	5–6	Boy	11.2 (8.6)	0.53 (0.42)
	5–6	Girl	10.6 (8.2)	0.51 (0.40)
Stifelman (2007)[b]	6	Boy	12.7 (9.8)	0.57[c]
	6	Girl	12.1 (9.3)	0.57[c]
This study	5–6	Both	10.1	0.51

[a] Values of daily inhalation rate ($m^3\,day^{-1}$) and normalized daily inhalation rate ($m^3\,kg\,day^{-1}$) in Brochu et al. and Stifelman are corrected with multiplying 1.3 in an assumption that the values in those studies are underestimated by 30% due to bias from the use of adult derived VQ values. Values in original article are in parentheses.
[b] Brochu et al. and Stifelman estimated daily inhalation rate by using physiologial approach using daily total energy expenditure measured with DLW method.
[c] Estimated by dividing the daily inhalation rate by the hypothetical body weight of 6-year-old boys and girls 22.1 and 21.3 kg, respectively.

While the strength of the approach we used in this study was to estimate the daily inhalation rate using an objective measure of physical activity, time–activity logs during periods when the accelerometer was not worn were still necessary to supplement this missing data. Our approach was also limited by the availability of $\dot{V}_E$ data in children. More data on $\dot{V}_E$ corresponding to a variety of physical activities observed during daily life in children will provide a more reliable estimate of their daily inhalation rate.

In conclusion, we obtained high between-day stability of the daily inhalation rate estimated from ActiveTracer measurements over 3 days, including one weekend day. However, more than 3 days are required to improve the between-day reliability of daily time spent in different levels of physical activity in preschool children when using ActivTracer. The daily inhalation rate of preschool children in the current study is possibly comparable to that estimated from the daily energy expenditure measured with the DLW method with consideration of potential bias from a physiological parameter derived from adults.

Acknowledgments

The authors thank all of the participants, their parents and teachers, and the administration of the kindergarten for their kind participation and support. We also wish to thank Mr. Suzuki T. at GMS Inc., Japan for technical advice on the operation and maintenance of the ActivTracer accelerometers.

References

Allan M, Richardson GM. Probability density functions describing 24-hour inhalation rates for use in human health risk assessments. Hum Ecol Risk Assess 1998;4(2):379–408.
Arcus-Arth A, Blaisdell RJ. Statistical distributions of daily breathing rates for narrow age groups of infants and children. Risk Anal 2007;27(1):97-110.
Bender JM, Brownson RC, Elliott MB, Haire-Joshu DL. Children's physical activity: using accelerometers to validate a parent proxy record. Med Sci Sports Exerc 2005;37(8):1409–13.

Bland JM, Altman DG. 1986 statistical method for assessing agreement between two methods of clinical measurement. Lancet 1986;1:308–10.
Brochu P, Ducre-Robitaille JF, Brodeur J. Physiological daily inhalation rates for free-living individuals aged 1 month to 96 years, using data from doubly labeled water measurements: a proposal for air quality criteria, standard calculations and health risk assessment. Hum Ecol Risk Assess 2006a;12(4):675–701.
Brochu P, Ducre-Robitaille JF, Brodeur J. Physiological daily inhalation rates for free-living individuals aged 2.6 months to 96 years based on doubly labeled water measurements: comparison with time–activity–ventilation and metabolic energy conversion estimates. Hum Ecol Risk Assess 2006b;12(4):736–61.
FAO. Human energy requirements. Report of a joint FAO/WHO/UNU expert consultation, Rome, 17–24 October 2001; 2004. p. 35–52.
ICRP (International Commission on Radiological Protection). Report of the task group on reference man. Tokyo Japan: Elsevier Science; 1975.
IOM (Institute of Medicine). Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids, cholesterol, protein, and amino acids (macronutrients). Washington, DC: Food and Nutrition Board, National Academy Press; 2005. p. 880–935.
Janz KF, Witt J, Mahoney LT. The stability of children's physical activity as measured by accelerometry and self-report. Med Sci Sports Exerc 1995;27:1326–32.
Kawahara J, Tanaka C, Tanaka S. Estimation of minute respiratory ventilation rate of preschool children using tri-axial accelerometer. J Jpn Soc Atmos Environ 2010;45 (5):235–45.
Kawahara J, Tanaka S, Tanaka C, Hikihara Y, Aoki Y, Yonemoto J. Estimation of respiratory ventilation rate of preschool children in daily life by using an accelerometer. J Air Waste Manage Assoc 2011;61:46–54.
Layton DW. Metabolically consistent breathing rates for use in dose assessments. Health Phys 1993;64(1):23–36.
Lichtman SW, Pisarska K, Berman ER, Pestone M, Dowling H, Weisel H, et al. Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. N Engl J Med 1992;327(27):1893–8.
MHLW (Ministry of Health and Labor Welfare, Japan). Pharmaceutical and Food Safety Bureau. Guideline for assessing health risk of pesticide used indoor environment (a proposal). http://www.env.go.jp/water/dojo/noyaku/hisan_risk/hyoka_tih/com08/mat03.pdf. 2007 November.
Puyau MR, Adolph AL, Vohra FA, Zakeri I, Butte NF. Prediction of activity energy expenditure using accelerometers in children. Med Sci Sports Exerc 2004;36(9): 1625–31.
Stifelman M. Using doubly-labeled water measurements of human energy expenditure to estimate inhalation rates. Sci Total Environ 2007;373(2–3):585–90.
Tanaka C, Tanaka S, Kawahara J, Midorikawa T. Tri-axial accelerometry for assessment of physical activity in young children. Obesity 2007;15(5):1233–41.
Taylor CM, Pye OF, Caldwell AB. The energy expenditure of 9-to-11-year-old boys and girls (1) standing, drawing and (2) dressing and undressing. J Nutr 1948;36:123–31.
Torun B. Inaccuracy of applying energy expenditure rates of adults to children. Am J Clin Nutr 1983;38(5):813–4.
Trost SG, Pate RR, Freedson PS, Sallis JF, Taylor WC. Using objective measures with youth: how many days of monitoring are needed? Med Sci Sports Exerc 2000;32:426–31.
U.S. EPA. Exposure factors handbook. Washington, DC: National Center for Environmental Assessment. Office of Research and Development; 1997.
U.S. EPA. Metabolically-derived human ventilation rates: a revised approach based upon oxygen consumption rates. External review draft. Washington, DC: Office of Research and Development. National Center for Environmental Assessment; 2006.
U.S. EPA. Child-specific exposure factors handbook. Washington, DC: National Center for Environmental Assessment. Office of Research and Development; 2008.
WHO. IPCS principle for evaluating health risks in children associated with exposure to chemicals. Environmental health criteria, 237. Geneva, Switzerland: WHO Press; 2006.
Yamamura C, Tanaka S, Futami J, Oka J, Ishikawa-Takata K, Kashiwazaki H. Activity diary method for predicting energy expenditure as evaluated by a whole body indirect human calorimeter. J Nutr Sci Vitaminol 2003;49:262–9.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 285, NO. 1, pp. 444–452, January 1, 2010

Structural and Dynamic Features of the MutT Protein in the Recognition of Nucleotides with the Mutagenic 8-Oxoguanine Base*[S]

Received for publication, September 15, 2009, and in revised form, October 14, 2009 Published, JBC Papers in Press, October 28, 2009, DOI 10.1074/jbc.M109.066373

Teruya Nakamura[‡], Sachiko Meshitsuka[§], Seiju Kitagawa[§], Nanase Abe[§], Junichi Yamada[§], Tetsuya Ishino[§], Hiroaki Nakano[§], Teruhisa Tsuzuki[¶], Takefumi Doi[§], Yuji Kobayashi[§], Satoshi Fujii[‖], Mutsuo Sekiguchi[**], and Yuriko Yamagata[‡1]

From the [‡]Graduate School of Pharmaceutical Sciences, Kumamoto University, Kumamoto 862-0973, the [§]Graduate School of Pharmaceutical Sciences, Osaka University, Suita 565-0871, the [¶]Graduate School of Medical Sciences, Kyushu University, Fukuoka 812-8582, the [‖]School of Pharmaceutical Sciences, University of Shizuoka, Shizuoka 422-8526, and the [**]Fukuoka Dental College, Fukuoka 814-0193, Japan

Escherichia coli **MutT hydrolyzes 8-oxo-dGTP to 8-oxo-dGMP, an event that can prevent the misincorporation of 8-oxoguanine opposite adenine in DNA. Of the several enzymes that recognize 8-oxoguanine, MutT exhibits high substrate specificity for 8-oxoguanine nucleotides; however, the structural basis for this specificity is unknown. The crystal structures of MutT in the apo and holo forms and in the binary and ternary forms complexed with the product 8-oxo-dGMP and 8-oxo-dGMP plus Mn^{2+}, respectively, were determined. MutT strictly recognizes the overall conformation of 8-oxo-dGMP through a number of hydrogen bonds. This recognition mode revealed that 8-oxoguanine nucleotides are discriminated from guanine nucleotides by not only the hydrogen bond between the N7-H and Oδ (N119) atoms but also by the *syn* glycosidic conformation that 8-oxoguanine nucleotides prefer. Nevertheless, these discrimination factors cannot by themselves explain the roughly 34,000-fold difference between the affinity of MutT for 8-oxo-dGMP and dGMP. When the binary complex of MutT with 8-oxo-dGMP is compared with the ligand-free form, ordering and considerable movement of the flexible loops surrounding 8-oxo-dGMP in the binary complex are observed. These results indicate that MutT specifically recognizes 8-oxoguanine nucleotides by the ligand-induced conformational change.**

Although spontaneous mutations are indispensable to the evolutionary process of living organisms, they can also be lethal to the organism. Among the various modified bases in DNA, RNA, and nucleotides, 8-oxoguanine (8-oxoG),[2] a damaged form of guanine (G) generated by reactive oxygen species, is known to have highly mutagenic potency because of its mispairing with adenine. Therefore, organisms have an error avoidance pathway for preventing mutations caused by 8-oxoG. The *Escherichia coli* MutT protein (129 amino acids, M_r = 14,900) hydrolyzes 8-oxo-dGTP and 8-oxo-GTP to their corresponding nucleoside monophosphates and inorganic pyrophosphate in the presence of Mg^{2+} (1, 2). Because 8-oxo-dGTP and 8-oxo-GTP can be misincorporated opposite adenine by DNA and RNA polymerases, the hydrolysis of the damaged nucleotides by MutT can avoid replicational and transcriptional errors. In DNA, 8-oxoG paired with cytosine is excised by MutM, an 8-oxoG DNA glycosylase, whereas MutY, an adenine DNA glycosylase, removes adenine paired with 8-oxoG (3–6).

The substrate specificities of enzymes that recognize 8-oxoG are quite varied. MutT exhibits high substrate specificity for 8-oxoG nucleotides; that is, the K_m for 8-oxo-dGTP is 14,000-fold lower than that for dGTP (7). In contrast, human MutT homologue 1 (hMTH1) hydrolyzes not only 8-oxo-dGTP but also several oxidized purine nucleotides such as 2-oxo-dATP, 2-oxo-ATP, 8-oxo-dATP, and 8-oxo-ATP. In terms of the hydrolysis of 8-oxo-dGTP, the K_m of hMTH1 for 8-oxo-dGTP is only 17-fold lower than that for dGTP (8, 9). The solution structure of hMTH1 as determined by NMR has revealed its overall architecture and possible substrate-binding region (10); however, the broad substrate recognition mechanism of hMTH1 remains to be elucidated. MutM and MutY also have low specificity for 8-oxoG. For example, MutM can recognize a variety of damaged bases such as formamidopyrimidine, 5-hydroxycytosine, and dihydrouracil in addition to 8-oxoG (11–13), and MutY shows a kinetic preference for A:8-oxoG that is only 6-fold greater than that for A:G (14).

The crystal structures of OGG1, MutM, and MutY complexed with 8-oxoG-containing DNA (13, 15, 16) have revealed

* This work was supported in part by grants-in-aid for scientific research and the National Project for Protein Structural and Functional Analysis from the Ministry of Education, Culture, Sports, Sciences and Technology of Japan.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S4.

The atomic coordinates and structure factors (codes 3A6S, 3A6T, 3A6U, and 3A6V) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] To whom correspondence should be addressed: Graduate School of Pharmaceutical Sciences, Kumamoto University, 5-1 Oe-honmachi, Kumamoto 862-0973, Japan. Tel./Fax: 81-96-371-4638; E-mail: yamagata@gpo.kumamoto-u.ac.jp.

[2] The abbreviations used are: 8-oxoG, 8-oxoguanine; Nudix, nucleoside diphosphate linked to some other moiety, X; SLHL, strand-loop-helix-loop; r.m.s.d., root mean square deviation; AMPCPP, adenosine 5′-(α,β-methyleneriphosphate); SeMet, selenomethionine; hMTH1, human MutT homologue 1.

that, interestingly, OGG1 and MutM do not recognize the O8 atom, which is the most characteristic feature of the 8-oxoG moiety, and the interaction observed between the O8 atom and the main-chain atom of MutY is relatively weak. Alternatively, OGG1, MutM, and MutY commonly discriminate 8-oxoG from G by the protonation at N7 accompanied by the oxidation of C8. Structural studies on various enzymes that recognize 8-oxoG have succeeded in explaining the mechanism by which 8-oxoG is discriminated from normal G in DNA, but one of the most interesting questions to be elucidated is the mechanism by which MutT acquires extremely high substrate specificity for 8-oxoG compared with the other enzymes.

MutT belongs to the Nudix (nucleoside diphosphate linked to some other moiety, X) hydrolase family (17). Nudix family members have a highly conserved MutT signature (Nudix motif); *i.e.* $GX_5EX_7REUXEEXGU$, where U is a hydrophobic residue and X is any amino acid. Current genome analyses have found a large number of open reading frames containing the MutT signature, but their functions, *i.e.* their substrates, are not identified in the case of almost all these proteins because of a lack of homology outside the MutT signature. MutT is the most examined protein in this family. Its structure was first determined by NMR (18) and has greatly contributed to the study of the Nudix hydrolase family. NMR studies of MutT with its product, 8-oxo-dGMP, have predicted several recognition models of 8-oxo-dGMP (19). However, the precise recognition mechanism of 8-oxoG nucleotides remains unclear. Therefore, it is necessary to determine the crystal structures of MutT to explain the extremely high substrate specificity of MutT for 8-oxoG nucleotides.

Here, we present x-ray crystallographic analyses of the apo enzyme; the Mn^{2+}-bound holo enzyme (MutT-Mn^{2+}); the binary complex with 8-oxo-dGMP, a reaction product (MutT-8-oxo-dGMP); and the tertiary complex with 8-oxo-dGMP and Mn^{2+} (MutT-8-oxo-dGMP-Mn^{2+}). These structures have revealed the mechanism of the extremely high substrate specificity of MutT for 8-oxoG nucleotides and have allowed us to propose the exact roles of some conserved residues in the MutT signature.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification—The *E. coli* strain BL21 (DE3) harboring a newly constructed pET8c/MutT plasmid was used for the expression of native and selenomethionine (SeMet)-substituted MutTs. Native MutT was overexpressed in Luria-Bertani (LB) broth, and SeMet MutT was overexpressed in LeMaster broth containing seleno-DL-methionine instead of methionine with sufficient amounts of isoleucine, lysine, and threonine to inhibit the methionine pathway (20, 21). This condition was also present in the overexpression of SeMet hMTH1 (22). Purification of MutT was carried out by almost the same procedure (except that the hydroxyapatite column chromatography step was skipped), as described previously (23). DEAE-Sepharose and the HiPrep 16/60 Sephacryl S-200 HR column were substituted for DEAE-Sephacel and the Sephadex G-75 column, respectively. The purified protein solution was concentrated to ~6 mg/ml.

Crystallization—The native and SeMet-substituted apo forms and all complexes were crystallized by hanging drop vapor diffusion at 288 K. Crystals of native and SeMet-substituted forms were obtained from a droplet containing 3 mg/ml protein, 10 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 2.5% glycerol, 0.5 mM 2-mercaptoethanol, 0.7 M potassium sodium tartrate, and 44 mM HEPES-NaOH (pH 7.5) equilibrated against a reservoir containing 1.4 M potassium sodium tartrate and 87 mM HEPES-NaOH (pH 7.5). Crystals of MutT-Mn^{2+} were obtained in the same manner, as described above, except that 10 mM $MnCl_2$ was added to the droplet. The crystallizations of MutT-8-oxo-dGMP and MutT-8-oxo-dGMP-Mn^{2+} were described previously (24). The crystals were transferred to a cryosolution of each reservoir containing 30% sucrose and were then flash frozen.

Data Collection, Processing, Phasing, and Structure Refinement—Diffraction data were collected at 100 K on beamline 18B of the Photon Factory (Tsukuba, Japan) and on beamlines 41XU, 44XU, 38B1, and 40B2 of SPring-8 (Harima, Japan). The data for native and SeMet derivative forms were processed and scaled by DSP/MOSFLM and SCALA (25). The data for MutT-Mn^{2+} were processed and scaled by DENZO and SCALEPACK (26). There are two molecules in the asymmetric unit with V_M of 2.2 (native MutT) and 2.5 (MutT-Mn^{2+}) $Å^3 \cdot Da^{-1}$ (27). Data collection statistics of the best data used for structure determination and refinements are listed in Table 1. Data collection statistics of MutT-8-oxo-dGMP and MutT-8-oxo-dGMP-Mn^{2+} are quoted from the reference by Nakamura *et al.* (24).

The positions of eight selenium atoms were determined using SOLVE (28). The initial phases were calculated using MLPHARE (29) and improved using DM (30). The initial model was built using TOM (31) and O (32). The model was refined using X-PLOR (33) and CNS (34). Using the model of the SeMet derivative, the successive refinement of native MutT converged at an R value of 20.4% and an R_{free} of 23.1% for reflections in the resolution range 20–1.8 Å. The structure of MutT-8-oxo-dGMP was solved by molecular replacement with AMoRe (35) using the structure of the native apo form as a search model. The $2F_o - F_c$ maps after CNS refinements clearly showed the density for 8-oxo-dGMP and the conformationally changed loop regions (L-A and L-D). These regions were manually built and fitted into the density with O. The structure of MutT-8-oxo-dGMP-Mn^{2+} was refined starting with the coordinates of the MutT-8-oxo-dGMP. The structure of MutT-Mn^{2+} was solved by molecular replacement with AMoRe by using the structure of the apo form as a search model. The stereochemical qualities of the structures were checked by PROCHECK (36); the refinement statistics are listed in Table 2. Superposition of MutT structures were carried out using Lsqkab (37). All molecular graphics were prepared using PyMOL (38).

RESULTS AND DISCUSSION

Overall Structures of MutT and MutT-8-oxo-dGMP—The crystal structures of the MutT apo and MutT-8-oxo-dGMP complex forms were determined at a resolution of 1.8 and 1.96 Å, respectively. MutT is composed of two α-helices (α-1 and α-2) and six β-strands (β-1 to β-6) (Figs. 1 and 2*A*); it adopts an

TABLE 1

Data collection statistics

Values in parentheses correspond to the highest resolution shell.

Diffraction data	MutT				MutT–8-oxo-dGMP	MutT–8-oxo-dGMP–Mn^{2+}
	Native	Peak	Edge	Remote		
Beam line	SPring-8 BL41XU		PF BL18B		SPring-8 BL41XU	SPring-8 BL40B2
Wavelength (Å)	0.7080	0.9793	0.9791	0.9500	0.9000	1.296
Space group	$P2_1$		$P2_1$		$P2_12_12_1$	$P2_12_12_1$
Unit-cell lengths (Å, °)	$a = 33.9$		$a = 34.1$		$a = 37.9$	$a = 38.2$
	$b = 71.6$		$b = 71.1$		$b = 56.0$	$b = 56.0$
	$c = 55.8$		$c = 55.7$		$c = 59.4$	$c = 59.3$
	$\beta = 99.0$		$\beta = 98.7$			
Resolution range (Å)	20.0–1.8 (1.9–1.8)		20.0–2.2 (2.3–2.2)		20.0–1.96 (2.08–1.96)	18.56–2.56 (2.72–2.56)
No. of observed reflections	85,618	49,308	49,226	50,094	54,195	26,502
No. of unique reflections	24,217	13,041	13,015	13,147	9,344	4,395
Completeness (%)	99.7 (99.7)	97.2 (97.2)	97.3 (97.3)	98.3 (98.3)	97.6 (93.2)	99.2 (95.7)
R_{merge} [a] (%)	3.1 (9.4)	3.9 (12.1)	3.8 (12.3)	3.9 (12.1)	6.5 (18.1)	7.7 (15.2)
$\langle I/\sigma \rangle$	15.8 (7.9)	8.8 (3.3)	9.4 (5.0)	8.8 (5.6)	29.6 (6.2)	45.9 (22.6)

	MutT-Mn^{2+}
Beam line	PF BL18B
Wavelength (Å)	1.000
Space group	$P2_1$
Unit-cell lengths (Å, °)	$a = 35.8$
	$b = 56.0$
	$c = 74.1$
	$\beta = 96.4$
Resolution range (Å)	40.0–2.0 (2.03–2.00)
No. of observed reflections	77,213
No. of unique reflections	19,810
Completeness (%)	98.1 (96.2)
R_{merge} [a] (%)	4.2 (6.4)
$\langle I/\sigma \rangle$	28.6 (18.6)

[a] $R_{merge} = 100 \times \Sigma|I_{hkl} - \langle I_{hkl} \rangle|/\Sigma I_{hkl}$, $\langle I_{hkl} \rangle$ is the mean value of I_{hkl}.

TABLE 2

Refinement statistics

Diffraction data	MutT	MutT–8-oxo-dGMP	MutT–8-oxo-dGMP–Mn^{2+}	MutT-Mn^{2+}
Resolution range (Å)	20.0–1.8	20.0–1.96	18.56–2.56	20.0–2.0
Number of reflections used	24,213	9,280	4,394	19,386
Number of atoms				
Protein	2,025	1,038	1,029	2,010
Water	190	136	92	131
Nucleotide	0	24	24	0
Mn^{2+} ion	0	0	1	4[a]
Other	11	34	5	32
Completeness (%)	99.1	97.3	99.9	98.1
R_{cryst}/R_{free} [b] (%)	20.4/23.1	17.8/20.1	19.3/24.2	19.2/22.7
Ramachandran plot (%)				
Most favored	93.0	91.5	87.6	94.9
Additional allowed	7.0	8.5	12.4	5.1
Generously allowed	0	0	0	0
Disallowed	0	0	0	0
r.m.s.d. in bonds (Å)	0.005	0.005	0.007	0.005
r.m.s.d. in angles (°)	1.2	1.3	1.3	1.2

[a] Two ions per monomer.

[b] $R_{cryst} = 100 \times \Sigma||F_o| - |F_c||/\Sigma|F_o|$. R_{free} was calculated from the test set (5% of the total data).

α-β-α sandwich structure that is conserved among members of the Nudix family. The Nudix motif (23 residues from Gly-38 *i.e.* the MutT signature ($GX_5EX_7REUXEEXGU$), adopts the characteristic strand-loop-helix-loop (SLHL) structure formed by β-3′, L-B, α-1, and L-C (39, 40). The crystal of the apo form contains two protein molecules per asymmetric unit, and they are very similar to each other with root mean square deviation (r.m.s.d.) of 0.5 Å for the corresponding 121 Cα atoms. For simplicity, only one molecule will be referred to in all further discussions. MutT exists as a monomer, which is found in the MutT-8-oxo-dGMP crystal.

In the apo form, the electron densities of L-A connecting β-2 and β-3 are not available, indicating that the L-A loop region has a highly flexible conformation (Fig. 2*A*). On the other hand, in the MutT-8-oxo-dGMP complex, the ordering of the flexible L-A loop by interactions with 8-oxo-dGMP was observed (Fig. 2, *B* and *C*). The plot of the displacement between the Cα atoms of the apo and complex forms is shown in supplemental Fig. S1*A*. The movements of the L-A and L-D regions are large (~8–10 Å) (Fig. 2*C* and supplemental Fig. S1*A*). Except for these loop regions, the two forms have a similar structure with an r.m.s.d. of 0.9 Å for the corresponding 101 Cα atoms.

A structural similarity search performed using the DALI server (41), with the coordinates of MutT-8-oxo-dGMP, indicated that 62 proteins (154 Protein Data Bank (PDB) ID numbers, 278 protein chains) are structural homologs of MutT with Z-scores of >6.0 and belong to the Nudix superfamily with the Nudix fold. The MutT structure, with two α-helices and six β-strands, comprises the smallest structural unit among members of the Nudix superfamily. Of 62 proteins, half have

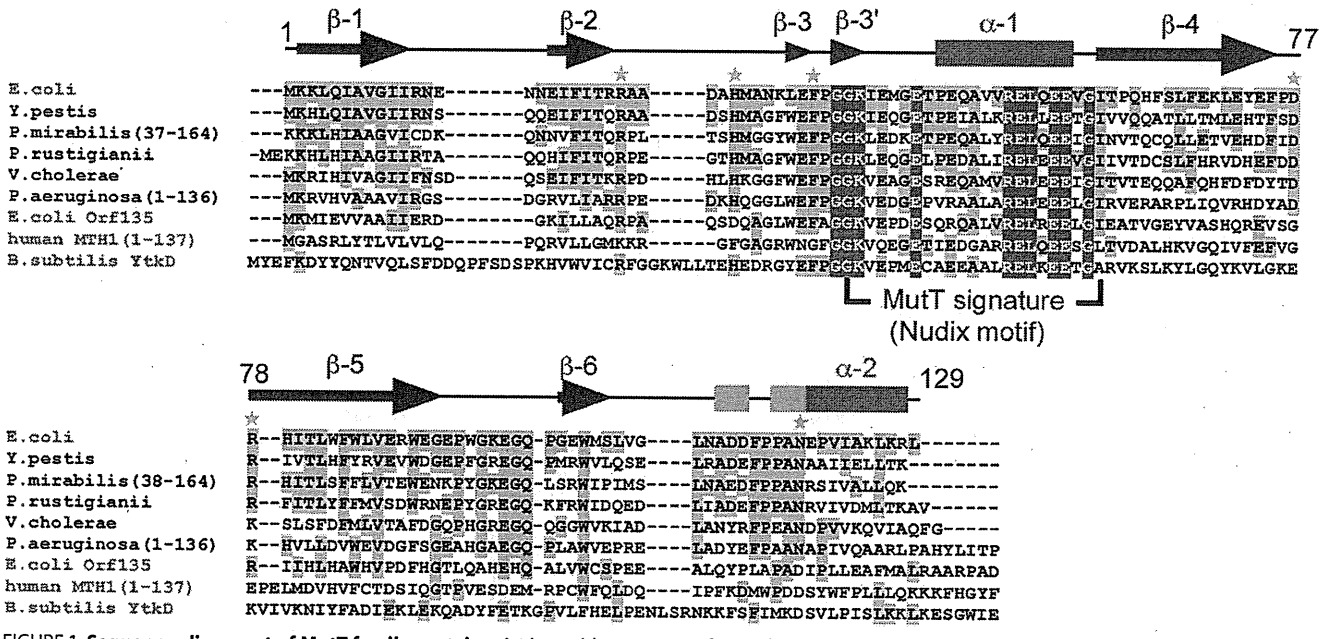

FIGURE 1. **Sequence alignment of MutT family proteins.** Amino acid sequences of MutT family proteins were aligned using ClustalW (72). MutT homologs from species related to *E. coli*, which share high sequence similarity, were chosen and are listed. They are from *E. coli* (CAA28523), *Yersinia pestis* (NP_670913), *Proteus mirabilis* (ZP_03840798), *Providencia rustigianii* (ZP_03315124), *Vibrio cholerae* (NP_232022), and *Pseudomonas aeruginosa* (ZP_04932260). In addition to *E. coli* MutT, *E. coli* Orf135 (BAA15549), human MTH1 (BAA07601), and *B. subtilis* YtkD (NP_390941), which have 8-oxo-dGTPase activity *in vitro*, were added and are shown in *red*. Absolutely conserved residues are shown in *red*, and identical residues are in *pink*. The *green asterisks* on the *E. coli* MutT sequence indicate amino residues that participate in the recognition of 8-oxoG and the ligand-induced conformational change. The secondary structure of *E. coli* MutT in the apo form is shown at the *top*. The α-helices, β-strands, and 3_{10} helices are represented as *red bars*, *blue arrows*, and *pink bars*, respectively.

unknown functions. Structures that are highly similar to the MutT complex form are the monomer structures of *Bdellovibrio bacteriovorus* RNA pyrophosphohydrolase; *i.e.* BdRppH in the ternary and binary forms complexed with GTP and Mg^{2+} (BdRppH-GTP-Mg^{2+}, 3FFU, r.m.s.d. = 1.8 Å, Z = 19.0) and with dGTP (BdRppH-dGTP, 3EF5, r.m.s.d. = 1.9 Å, Z = 18.3) (40) and unknown proteins from *Bartonella henselae* (3HHJ, r.m.s.d. = 1.8 Å, Z = 19.8) and *Methanosarcina mazei* (3GRN, r.m.s.d. = 2.1 Å, Z = 17.5), respectively. In MutT, the r.m.s.d. is rather large: 3.3 Å for 120 Cα atoms between the x-ray and NMR structures in the ligand-free form and 3.5 Å for 127 Cα atoms between structures in the complex form (PDB IDs: 1MUT and 1PUS) (18, 19).

Recognition Scheme of 8-oxo-dGMP by MutT—When 8-oxo-dGMP binds to MutT, large ligand-induced conformational changes occur in the L-A and L-D regions, namely, the ordering of the flexible L-A loop and considerable movement of L-A and L-D to the surrounding 8-oxo-dGMP (Figs. 2*C* and 3*A*). The side chains of Arg-23 and His-28 on L-A form hydrogen-bonding interactions with phosphate and sugar moieties of 8-oxo-dGMP, respectively, whereas Arg-78 interacts with the sugar moiety through a water-mediated hydrogen bond. Asp-77 and Arg-78 make two hydrogen bonds between their side chains. The conformational change of the L-A and L-D regions also produces the water molecule-mediated interaction between His-28 and Asp-77 and the CH-π interaction between His-28 and Phe-75 (Fig. 3*A* and supplemental Fig. S2). Thus, the loops L-A and L-D connect to each other, resulting in the formation of a cave composed of β-1, β-3, β-3′, β-5, and α-2 for substrate binding (Fig. 3*B*). 8-Oxo-dGMP is inserted deeply into the cave in which the wall on one side is filled with hydrophobic residues (Leu-4, Ile-6, Val-8, Ile-80, and Leu-82), and the other side and bottom include some polar residues (Arg-23 and Asn-119). The 8-oxoG base and the deoxyribose are perfectly buried, and the phosphate group faces the solvents (Fig. 3*B*). The glycosidic conformation of 8-oxo-dGMP bound to MutT is *syn*. This fact is consistent with the first suggestion by Bessman *et al.* that MutT may recognize the *syn* conformation, because the 8-substituted purine nucleotides were better substrates compared with the normal purine nucleotides (42). The sugar ring puckering and the sugar-phosphate backbone conformation of 8-oxo-dGMP are C2′-*endo* and *gauche*$^-$-*trans*, respectively. These conformations are generally observed in 8-substituted purine nucleosides and 5′-nucleotides (43).

These ligand-induced conformational changes result in the strict recognition of the overall structure of 8-oxo-dGMP by MutT through a number of hydrogen bonds (Fig. 3*C*). The characteristic features of the 8-oxoG base are the oxygen atom (O8) at C8 and the hydrogen atom (N7-H) at N7 accompanied by oxidation. MutT recognizes N7-H of 8-oxoG by a hydrogen bond with Oδ of Asn-119 (Fig. 3*C*, a *red dashed line*). The 8-oxoG base is also recognized by hydrogen bonds with Asn-119 and Phe-35; *i.e.* the Nδ of Asn-119 forms a hydrogen bond with O6 of 8-oxoG, and the main-chain atoms of Phe-35 participate in three types of hydrogen bonds with N2-H, N1-H, and O6. On the other hand, the O8 atom does not form hydrogen bonds with any amino acid residues, although it does participate in the weak C-H–O interaction with the phenyl ring C-H of Phe-75 (C–O distance, 3.4 Å) and the van der Waals interactions with the side-chain C-H moieties of Phe-75, Pro-116, Leu-

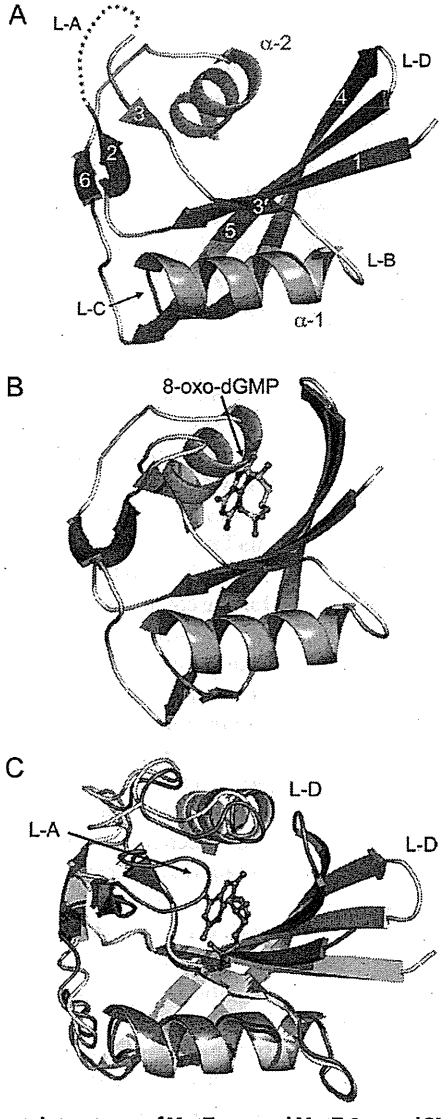

FIGURE 2. **Crystal structures of MutT apo and MutT-8-oxo-dGMP complex forms.** *A*, overall structure of MutT. α-Helices are in *pink*, and β-strands are in *slate*. A missing region of L-A is shown as a *gray dashed line*. *B*, overall structure of MutT-8-oxo-dGMP. 8-Oxo-dGMP is shown in *ball and stick representation*. *C*, comparison of the structures of the apo and complex forms. Apo and complex forms are shown in *gray* and *slate*, respectively. L-A and L-D regions in MutT-8-oxo-dGMP adopt a closed conformation as compared with those in the apo form.

82, Ile-80, and Ala-118 (Fig. 3*D*). In addition, the carbonyl oxygen of Gly-37 forms water molecule-mediated hydrogen bonds with N2-H of 8-oxoG and the phosphate oxygen. The side chain of His-28 is directly hydrogen-bonded to the O3′ atom of the deoxyribose. The O3′ atom also forms a hydrogen bond with a water molecule, binding to the side chain of Arg-78. The phosphate group forms a hydrogen bond with the side chain of Arg-23 and a water molecule-mediated hydrogen bond with the main chain of Lys-39. In summary, 8-oxo-dGMP is surrounded by 12 types of hydrogen bonds. The hydrogen-bonding interactions with the pyrimidine moiety and the α-phosphate group in MutT-8-oxo-dGMP are similar to those with the corresponding pyrimidine moieties and α-phosphate groups in the structures of BdRppH-GTP-Mg^{2+} and BdRppH-dGTP (BdRppH-(d)GTPs) (40). The positions of the base moieties of (d)GTPs with the *syn* conformation in BdRppH-(d)GTPs accord with that of 8-oxo-dGMP in MutT-8-oxo-dGMP with an r.m.s.d. of 0.6 Å for the corresponding 11 atoms when proteins are superimposed. BdRppH with Arg-40, Phe-52, and Asn-136 residues corresponding to Arg-23, Phe-35, and Asn-119 of MutT, respectively, recognizes N1-H, N2-H, and O6 of the pyrimidine moiety by four hydrogen bonds with Phe-52 and Asn-136; Pα-O of the α-phosphate group is recognized by a hydrogen bond with Arg-40. This recognition mode is the same as that observed in MutT. Apart from the similarities, differences are found in recognition of the imidazole moiety of the base and the sugar moiety as well as in the ligand-induced conformational change. The imidazole and sugar moieties of (d)GTPs in BdRppH-(d)GTPs do not form hydrogen bonds with any residues in BdRppH. In addition, although ligand-induced conformational change with loop ordering is observed in BdRppH, the change is significantly small (~2–4 Å) as compared with that in MutT. The large ligand-induced conformational change observed in MutT does contribute to its high affinity for 8-oxoG nucleotides, as discussed below. There is a large discrepancy between the K_m values of 0.081 and 268 μM for the hydrolysis of 8-oxo-dGTP by MutT and of dGTP by BdRppH, respectively (7, 44). This may derive from these structural differences, in addition to the unfavorable *syn* conformation of (d)GTPs in BdRppH-(d)GTPs.

As a result of the strict recognition of 8-oxo-dGMP with the large conformational change, there are low *B* factors and unambiguous electron densities around 8-oxo-dGMP (supplemental Fig. S1*B* and Fig. 3*E*). The average *B* factor of 8-oxo-dGMP is 12.4 $Å^2$ and that of the residues involved in the recognition of 8-oxo-dGMP, Arg-23, His-28, Phe-35, Asp-77, Arg-78, and Asn-119 is 13.6 $Å^2$. These low *B* factors and unambiguous electron densities represent the small thermal motion and/or the ordered positioning of 8-oxo-dGMP and the residues of the active site in the crystal lattice. This phenomenon explains isothermal titration calorimetry experiments (45), indicating that the tight binding of 8-oxo-dGMP to MutT ($\Delta G = -9.8$ kcal/mol) is driven by a highly favorable enthalpy ($\Delta H = -39.0$ kcal/mol) with an unfavorable entropy ($-T\Delta S = 29.2$ kcal/mol). The unfavorable entropy would be a result of the conformational rigidity generated from the connection of loops L-A and L-D with large ligand-induced conformational changes. On the other hand, the more favorable enthalpy would be produced by the large number of hydrogen bonds and van der Waals interactions formed between 8-oxo-dGMP and MutT; this is sufficient to compensate for the unfavorable entropy and to bind tightly.

Furthermore, the hydrogen bond-mediated recognition mode found in MutT-8oxo-dGMP is comparable with the results of mutational studies in which it was found that the R78A, N119D, and N119A mutants show 7-, 37-, and 1650-fold decreases in affinity for 8-oxo-dGMP in comparison with the wild type and that they lose binding free energies ($\Delta\Delta G$) of 1.1, 2.1, and 4.3 kcal/mol, respectively, as measured by the increases in K_I (46). According to previous reports, the contribution of the hydrogen bond to protein stability can be estimated as ~2 and 1.2 kcal/mol for hydrogen bonds between protein residues

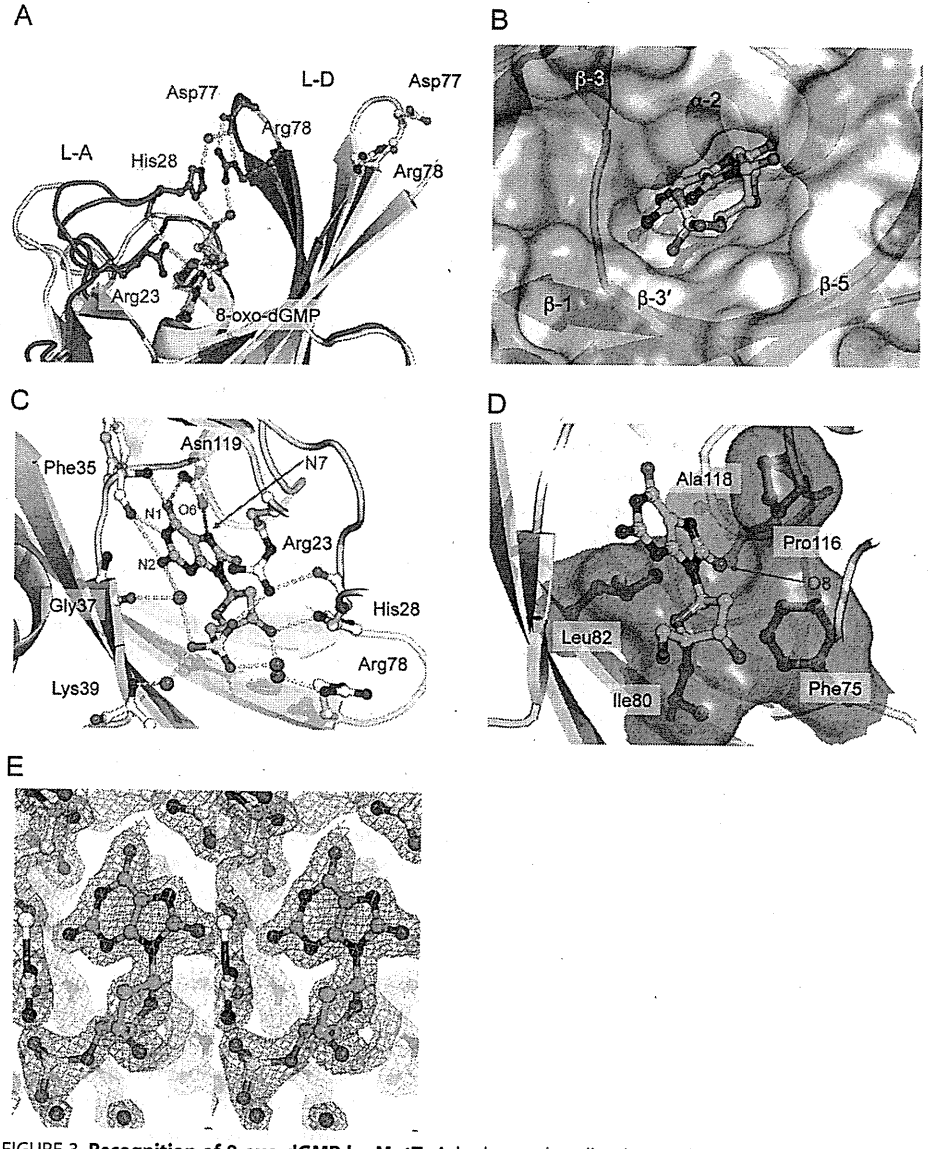

FIGURE 3. **Recognition of 8-oxo-dGMP by MutT.** *A*, hydrogen bonding interactions between 8-oxo-dGMP and loop regions (apo in *gray* and MutT-8-oxo-dGMP in *slate*). Amino acid residues involved in the hydrogen bonding interactions are shown in *ball and stick representation*. Water molecules are in *red*. Hydrogen bonds are shown as *yellow dashed lines*. *B*, the hydrophobic cave composed of β-1, β-3, β-3′, β-5, and α-2 is represented as a *translucent surface* (carbon in *white*, nitrogen in *cyan*, and oxygen in *pink*). *C*, interactions for the *syn* conformation of 8-oxo-dGMP. The hydrogen bond between Oδ of Asn-119 and N7-H of 8-oxoG is shown as a *red dashed line*. *D*, van der Waals interactions around the O8 atom. Amino acid residues recognizing O8 are shown in *ball, stick*, and *translucent surface*. *E*, a $2F_o - F_c$ electron density map around 8-oxo-dGMP contoured at 1.5 σ (*stereo view*).

and for those between a water molecule and a protein residue, respectively (47, 48). Judging from the MutT-8oxo-dGMP structure, the N119D and N119A mutants lose one hydrogen bond between the O6 of 8-oxo-dGMP and the Nδ-H of Asn-119 and two hydrogen bonds involving the O6 and N7-H of 8-oxo-dGMP and the amide group of Asn-119, respectively (Fig. 3*C*). The R78A mutant loses the hydrogen bond to a water molecule (Fig. 3*C*); that is, $\Delta\Delta G$ losses of 1.2, 2, and 4 kcal/mol are estimated for the R78A, N119D, and N119A mutants, respectively; this agrees perfectly with the experimental data (46).

The feature of the substrate-binding site in MutT-8-oxo-dGMP is also consistent with reports that MutT hydrolyzes both deoxyribose and ribose derivatives of 8-oxoG nucleotides with similar efficiency (2, 7). This is because, despite a number of hydrogen-bonding interactions between 8-oxo-dGMP and MutT, there is space for a hydroxyl group instead of the hydrogen atom at the 2′ position of the sugar ring (Fig. 3*B*). This recognition mechanism of 8-oxo-dGMP by MutT is different from any models predicted by NMR studies (PDB IDs: 1PPX, 1PUN, 1PUQ, 1PUS, and 1MUT) (supplemental Fig. S3).

Discrimination of 8-oxoG Nucleotides from G Nucleotides—The K_m values of *E. coli* MutT for 8-oxo-dGTP and 8-oxo-GTP are ~3,800- to 14,000-fold lower than the values for the corresponding G nucleotides (7). These data agree with the observation that the K_d value (52 nM) between 8-oxo-dGMP and MutT is 34,000-fold lower than that (1.76 mM) between dGMP and MutT (45). Thus, the most important question that this study addresses is the mechanism by which MutT obtains high substrate specificity for 8-oxoG nucleotides as compared with G nucleotides.

According to the recognition scheme of 8-oxo-dGMP by MutT, the major difference in the recognition of 8-oxoG *versus* G is whether the single hydrogen bond between Oδ of Asn-119 and N7-H of 8-oxoG occurs or not. This situation is similar to those of OGG1, MutM, and MutY (13, 15, 16). If the side-chain conformation of Asn-119 in the MutT complex with G nucleotides was the same as that in the MutT-8-oxo-dGMP structure, the two lone pairs at N7 of G and Oδ of Asn-119 would be repulsive (supplemental Fig. S4). To avoid this repulsion, a rotation about the side-chain torsion angles in Asn-119 should be required. Thus, the difference in the number of hydrogen bonds formed between MutT-G and MutT-8-oxoG complexes is only one. The contribution of one hydrogen bond to $\Delta\Delta G$ is estimated to be 2 kcal/mol (47, 48).

The *syn* glycosidic conformation of 8-oxo-dGMP must also be one of the elements contributing to the substrate specificity of MutT, because 8-oxoG nucleotides favor a *syn* conformation by the steric hindrance between O8 and the sugar moiety; this is in contrast with G nucleotides that adopt both *syn* and *anti* conformations (49, 50). Because of the lack of quantitative data

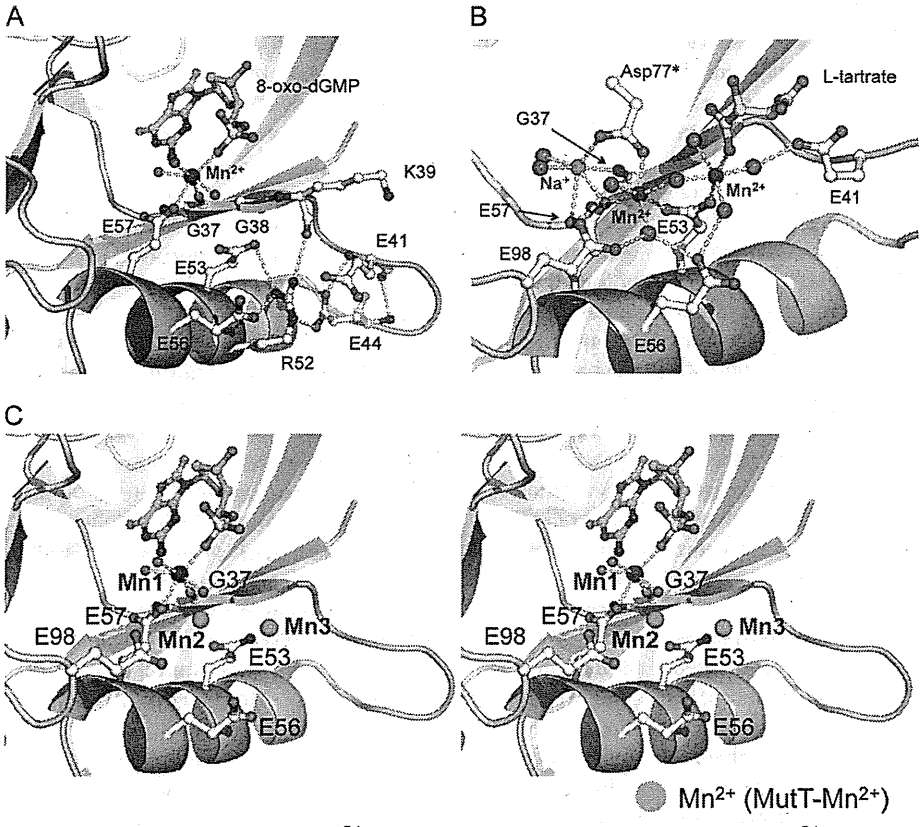

FIGURE 4. **Coordination scheme of Mn^{2+} at the MutT signature in MutT-8-oxo-dGMP-Mn^{2+} and MutT-Mn^{2+}.** *A*, the coordination scheme of Mn^{2+} and the structure of the MutT signature in MutT-8-oxo-dGMP-Mn^{2+}. Mn^{2+} in *blue* has an ideal octahedral coordination with Gly-37, Glu-57, Pα-O of 8-oxo-dGMP, and water molecules. The hydrogen bonding interactions shown in *yellow dashed lines* contribute to the conformational stabilization of the SLHL structure of the MutT signature. The SLHL structure is in *pink*. *B*, the coordination scheme of Mn^{2+} in MutT-Mn^{2+}. Na^+ is shown in *green*. Asp-77* is an amino acid of another molecule in the asymmetric unit. *C*, superposition of two Mn^{2+} ions in MutT-Mn^{2+} onto the structure of MutT-8-oxo-dGMP-Mn^{2+} (*stereo view*). *Coloring* is as in *A*. Mn^{2+} ions observed in MutT-Mn^{2+} are shown in *cyan* (*Mn2* and *Mn3*).

on the preference of the *syn* conformation about the 8-oxoG nucleotides, it is difficult to estimate the free energy difference between the *syn* and *anti* conformations. However, it is unlikely that the preference of MutT for the *syn* conformation over the *anti* conformation is more than 10-fold ($\Delta\Delta G \approx 1.4$ kcal/mol), because the *anti* conformation for 8-oxoG nucleotides is sometimes observed in the crystal structures of 8-oxoG recognition complexes such as OGG1 and MutY (15, 16).

These discrimination factors cannot by themselves explain the roughly 34,000-fold difference between the binding affinity of MutT for 8-oxo-dGMP and dGMP ($\Delta\Delta G \approx 6$ kcal/mol). When 8-oxo-dGMP binds to MutT, large ligand-induced conformational changes with an ordering of loop regions are observed. On the other hand, in the binding of dGMP to MutT, the thermal parameters were $\Delta G = -3.7$, $\Delta H = -3.3$, and $-T\Delta S = -0.4$ kcal/mol; the changes in backbone ^{15}N and NH chemical shifts in 22 residues; and the slowing down of the NH exchange with D_2O of 20 residues are remarkably different when compared with the changes (the thermal parameters of $\Delta G = -9.8$, $\Delta H = -39.0$, and $-T\Delta S = 29.2$ kcal/mol; the changes in backbone ^{15}N and NH chemical shifts in 62 residues; and the slowing down of the NH exchange with D_2O of 45 residues) involved in the binding of 8-oxo-dGMP to MutT (19, 45). These facts suggest that no significant conformational change in MutT is observed when dGMP binds to MutT. The large ligand-induced conformational change in MutT also contributes to the discrimination of 8-oxoG nucleotides from G nucleotides.

A comparison of the amino acid sequences of MutT-related enzymes suggests that the enzymes with higher substrate specificity for 8-oxoG are only MutT homologs from closely related species with conserved amino acids in the positions that participate in the recognition of 8-oxoG and the conformational change (Fig. 1, *green asterisk*). In fact, these amino acids are not highly conserved among *E. coli* Orf135 (51), *Bacillus subtilis* YtkD (52), and hMTH1 (8) that have broad substrate specificities.

Structure of the MutT Signature and Metal-binding Sites—We have solved two types of Mn^{2+}-bound structures, MutT-8-oxo-dGMP-Mn^{2+} and MutT-Mn^{2+}, to determine metal-binding sites at the MutT signature of MutT. The crystal of MutT-Mn^{2+} contains two proteins per asymmetric unit. Their overall structures are very similar, with an r.m.s.d. of 0.3 Å for the corresponding 118 Cα atoms; for simplicity, only one molecule will be referred to in all further discussion. The structure of MutT-Mn^{2+} is similar to that of the apo form, with an r.m.s.d. of 0.6 Å for the corresponding 118 Cα atoms.

In the MutT signature having an SLHL structure (Fig. 4*A*), Gly-38, Glu-44, Arg-52, Glu-53, Glu-56, Glu-57, and Gly-59 are completely conserved among the members of the Nudix family (Fig. 1). The SLHL structure of MutT is similar to those of other enzymes in the Nudix family. For example, the conserved 23 residues can be superimposed on those found in the *Pyrobaculum aerophilum* Nudix protein with an r.m.s.d. of 0.5 Å (53). A characteristic feature in the SLHL structure is the hydrogen-bonding network centering on the converged Arg-52 that anchors α-1 to its connecting loop. The side chains of Glu-44 and Arg-52, which form two hydrogen bonds with each other, participate in hydrogen bonding to the main chain atoms in the nonconserved Glu-41 and Lys-39 residues, respectively. Arg-52 also interacts with the side chains of the conserved Glu-53 and Glu-56 residues (Fig. 4*A*). This hydrogen-bonding network is roughly the same as that in the other Nudix proteins and contributes to the conformational stability of the SLHL structure (39, 40, 53–65).

In the electron density maps of the MutT-Mn^{2+} crystal produced by co-crystallization, there were three peaks corresponding to the metal ions near the MutT signature. Judging from the peak heights, *B* factors, bond lengths, and bond angles, we determined that two peaks were Mn^{2+} ions (nearly ideal octahedral coordination and an average bond length of ~2.2 Å) and one was Na^+ (distorted octahedral coordination and an average bond length of ~2.5 Å) (Fig. 4*B*). Furthermore, the two Mn^{2+} sites were also confirmed from significant densities (>10 σ level) on the anomalous difference Fourier map at $\lambda = 1$ Å (data not shown). On the other hand, the densities of the Na^+ site were less than noise level. The refined MutT-Mn^{2+} structure reveals that two Mn^{2+} ions form a binuclear metal center with a bridged water molecule. One Mn^{2+} coordinates to the oxygen atoms of Glu-53, Glu-57, three water molecules, and Asp-77 in another molecule (Asp-77*). The coordination partners of another Mn^{2+} are Glu-53, L-tartrate (a crystallization reagent), and four water molecules. Na^+ coordinates to the oxygen atoms of Gly-37, Glu-57, Asp-77*, and two water molecules. Asp-77* and L-tartrate bind to the MutT signature through metal ions, but do not distort the SLHL structure.

In the MutT-8-oxo-dGMP-Mn^{2+} crystal (prepared by soaking MutT-8-oxo-dGMP crystals in reservoir supplemented with 1 mM $MnCl_2$), Mn^{2+} binds to the six oxygen atoms of the main chain of Gly-37, the side chain of Glu-57, the phosphate group, and three water molecules with nearly ideal octahedral coordination (Fig. 4*A*). The binding of Mn^{2+} to the MutT-8-oxo-dGMP binary complex makes the phosphate group move slightly toward Mn^{2+} (the P atom moves by 0.7 Å). The position of Mn^{2+} observed in MutT-8-oxo-dGMP-Mn^{2+} is close to that of Na^+ in MutT-Mn^{2+} (at a distance of 1.4 Å). The three Mn^{2+} binding sites consisting of the Mn^{2+} in MutT-8-oxo-dGMP-Mn^{2+} and two additional Mn^{2+} ions in MutT-Mn^{2+} (Fig. 4*C*) are located at common metal-binding sites observed in the Nudix family (60) and correspond to those observed in the ternary complexes of *E. coli* ADP-ribose pyrophosphatase, *Mycobacterium tuberculosis* ADP-ribose pyrophosphatase, *Caenorhabditis elegans* diadenosine 5′,5‴-P^1,P^4-tetraphosphate pyrophosphohydrolase, *Xenopus laevis* X29, human NUDT5, *Thermus thermophilus* Ndx2, and BdRppH (40, 55, 57, 64–67). Thus, the three sites are considered to be candidates for metal binding in 8-oxo-dGTP hydrolysis. Previous kinetic studies have shown that MutT binds to one Mn^{2+} in the absence of nucleotides and two Mn^{2+} ions in the presence of a nonspecific substrate analog, AMPCPP (68). The middle Mn^{2+} (*Mn2* in Fig. 4*C*) may be prebound to the active site, judging from the number of coordination partners in MutT and the Mn^{2+} ion (*Mn1* in Fig. 4*C*) found in MutT-8-oxo-dGMP-Mn^{2+}; otherwise, it and other Mn^{2+} ion (*Mn1* and *Mn3*, respectively in Fig. 4*C*) would be recruited with the substrate. The probability of the three metals binding to MutT in the presence of the real substrate 8-oxo-dGTP cannot be neglected because of the fact that the number of binding metals depends on the kinds of substrate analogs (40, 54, 60, 61, 64–66).

The structures of MutT-8-oxo-dGMP-Mn^{2+} and MutT-Mn^{2+} suggest structural insights into some essential or important residues for the 8-oxo-dGTP hydrolysis. Glu-53 and Glu-57 are essential for the suppression of spontaneous A:T to C:G transversion mutations (69), and E53Q and E57Q mutants decrease k_{cat} by 10^4 to 10^5-fold (70). On the other hand, Glu-56 is nonessential for the suppression of the mutations (69), and E56Q and E98Q mutants have relatively small effects (<24-fold) on k_{cat} (70). These results agree with our structural studies showing that essential residues, Glu-53 and Glu-57, directly bind to metal ions, whereas important residues, Glu-56 and Glu98, make water-mediated interactions with metal ions (Fig. 4, *A* and *B*). Gly-37 and Gly-38, which are located at the surface of the ligand-binding site, are also revealed to be essential residues for the suppression of the mutations (69). The side chains of any residues except Gly in positions 37 and 38 would contact the base moiety of the nucleotide ligand and the essential residue for the catalysis, Glu-53, respectively (Fig. 4, *A* and *C*). For this reason, to express 8-oxo-dGTPase activity, residues 37 and 38 must be Gly, which has the smallest side chain.

A number of kinetic, mutational, and NMR studies of MutT using dGTP and/or a substrate analog, AMPCPP, have been reported, and a catalytic mechanism is proposed by Mildvan and coworkers (71). Compared with our structural data, there appear to be some differences in the metal-binding sites. Gly-38 is involved in metal coordination in their model, but Gly-37, instead of Gly-38, is a metal ligand in the structure of MutT-8-oxo-dGMP-Mn^{2+} (Fig. 4, *A* and *C*). The all-crystal structures of Nudix proteins show that the carbonyl oxygen of the corresponding residue to Gly-38 participates in the formation of β-sheet, whereas that of Gly-37 binds to a metal ion (40, 55, 57–59, 61, 64–67). Glu-56 and Glu98, which are metal ligands in the model proposed by Mildvan and coworkers, interact with water molecules bound to metal ions in our structures. The metal coordination scheme changes in the active site during the reaction, and a more proper enzymatic mechanism activated by metal ions might be examined by kinetic protein crystallography.

REFERENCES

1. Maki, H., and Sekiguchi, M. (1992) *Nature* **355,** 273–275
2. Taddei, F., Hayakawa, H., Bouton, M., Cirinesi, A., Matic, I., Sekiguchi, M., and Radman, M. (1997) *Science* **278,** 128–130
3. Au, K. G., Cabrera, M., Miller, J. H., and Modrich, P. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85,** 9163–9166
4. Cabrera, M., Nghiem, Y., and Miller, J. H. (1988) *J. Bacteriol.* **170,** 5405–5407
5. Michaels, M. L., Cruz, C., Grollman, A. P., and Miller, J. H. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89,** 7022–7025
6. Tchou, J., Kasai, H., Shibutani, S., Chung, M. H., Laval, J., Grollman, A. P., and Nishimura, S. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88,** 4690–4694
7. Ito, R., Hayakawa, H., Sekiguchi, M., and Ishibashi, T. (2005) *Biochemistry* **44,** 6670–6674
8. Fujikawa, K., Kamiya, H., Yakushiji, H., Fujii, Y., Nakabeppu, Y., and Kasai, H. (1999) *J. Biol. Chem.* **274,** 18201–18205
9. Fujikawa, K., Kamiya, H., Yakushiji, H., Nakabeppu, Y., and Kasai, H. (2001) *Nucleic Acids Res.* **29,** 449–454
10. Mishima, M., Sakai, Y., Itoh, N., Kamiya, H., Furuichi, M., Takahashi, M., Yamagata, Y., Iwai, S., Nakabeppu, Y., and Shirakawa, M. (2004) *J. Biol. Chem.* **279,** 33806–33815
11. Tchou, J., Bodepudi, V., Shibutani, S., Antoshechkin, I., Miller, J., Grollman, A. P., and Johnson, F. (1994) *J. Biol. Chem.* **269,** 15318–15324
12. Hatahet, Z., Kow, Y. W., Purmal, A. A., Cunningham, R. P., and Wallace, S. S. (1994) *J. Biol. Chem.* **269,** 18814–18820
13. Fromme, J. C., and Verdine, G. L. (2003) *J. Biol. Chem.* **278,** 51543–51548
14. Porello, S. L., Leyes, A. E., and David, S. S. (1998) *Biochemistry* **37,**

14756–14764
15. Bruner, S. D., Norman, D. P., and Verdine, G. L. (2000) *Nature* **403,** 859–866
16. Fromme, J. C., Banerjee, A., Huang, S. J., and Verdine, G. L. (2004) *Nature* **427,** 652–656
17. Bessman, M. J., Frick, D. N., and O'Handley, S. F. (1996) *J. Biol. Chem.* **271,** 25059–25062
18. Abeygunawardana, C., Weber, D. J., Gittis, A. G., Frick, D. N., Lin, J., Miller, A. F., Bessman, M. J., and Mildvan, A. S. (1995) *Biochemistry* **34,** 14997–15005
19. Massiah, M. A., Saraswat, V., Azurmendi, H. F., and Mildvan, A. S. (2003) *Biochemistry* **42,** 10140–10154
20. Doublié, S. (1997) *Methods Enzymol.* **276,** 523–530
21. LeMaster, D. M., and Richards, F. M. (1985) *Biochemistry* **24,** 7263–7268
22. Nakamura, T., Kitaguchi, Y., Miyazawa, M., Kamiya, H., Toma, S., Ikemizu, S., Shirakawa, M., Nakabeppu, Y., and Yamagata, Y. (2006) *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **62,** 1283–1285
23. Akiyama, M., Maki, H., Sekiguchi, M., and Horiuchi, T. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86,** 3949–3952
24. Nakamura, T., Doi, T., Sekiguchi, M., and Yamagata, Y. (2004) *Acta Crystallogr. D Biol. Crystallogr.* **60,** 1641–1643
25. Rossmann, M. G., and van Beek, C. G. (1999) *Acta Crystallogr. D Biol. Crystallogr.* **55,** 1631–1640
26. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
27. Matthews, B. W. (1968) *J. Mol. Biol.* **33,** 491–497
28. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. D Biol. Crystallogr.* **55,** 849–861
29. Collaborative Computational Project, Number 4. (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50,** 760–763
30. Cowtan, K. (1994) *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **31,** 34–38
31. Cambillau, C., Horjales, E., and Jones, T. A. (1984) *J. Mol. Graph.* **2,** 53–54
32. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
33. Brünger, A. T. (1992) *A System for X-Ray Crystallography and NMR,* Yale University Press, New Haven, CT
34. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D Biol. Crystallogr.* **54,** 905–921
35. Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50,** 157–163
36. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
37. Kabsch, W. (1976) *Acta Crystallogr. Sect. A* **32,** 922–923
38. DeLano, W. L. (2002) *The PyMOL Molecular Graphics System,* DeLano Scientific LLC, San Carlos, CA
39. Gabelli, S. B., Bianchet, M. A., Xu, W., Dunn, C. A., Niu, Z. D., Amzel, L. M., and Bessman, M. J. (2007) *Structure* **15,** 1014–1022
40. Messing, S. A., Gabelli, S. B., Liu, Q., Celesnik, H., Belasco, J. G., Piñeiro, S. A., and Amzel, L. M. (2009) *Structure* **17,** 472–481
41. Holm, L., Kääriäinen, S., Rosenström, P., and Schenkel, A. (2008) *Bioinformatics* **24,** 2780–2781
42. Bhatnagar, S. K., Bullions, L. C., and Bessman, M. J. (1991) *J. Biol. Chem.* **266,** 9050–9054
43. Saenger, W. (1984) *Principles of Nucleic Acid Structure,* Springer-Verlag, New York
44. Steyert, S. R., Messing, S. A., Amzel, L. M., Gabelli, S. B., and Piñeiro, S. A. (2008) *J. Bacteriol.* **190,** 8215–8219
45. Saraswat, V., Massiah, M. A., Lopez, G., Amzel, L. M., and Mildvan, A. S. (2002) *Biochemistry* **41,** 15566–15577
46. Saraswat, V., Azurmendi, H. F., and Mildvan, A. S. (2004) *Biochemistry* **43,** 3404–3414
47. Funahashi, J., Takano, K., Yamagata, Y., and Yutani, K. (2002) *J. Biol. Chem.* **277,** 21792–21800
48. Takano, K., Yamagata, Y., Funahashi, J., Hioki, Y., Kuramitsu, S., and Yutani, K. (1999) *Biochemistry* **38,** 12698–12708
49. Kouchakdjian, M., Bodepudi, V., Shibutani, S., Eisenberg, M., Johnson, F., Grollman, A. P., and Patel, D. J. (1991) *Biochemistry* **30,** 1403–1412
50. Uesugi, S., Yano, J., Yano, E., and Ikehara, M. (1977) *J. Am. Chem. Soc.* **99,** 2313–2323
51. Kamiya, H., Murata-Kamiya, N., Iida, E., and Harashima, H. (2001) *Biochem. Biophys. Res. Commun.* **288,** 499–502
52. Xu, W., Jones, C. R., Dunn, C. A., and Bessman, M. J. (2004) *J. Bacteriol.* **186,** 8380–8384
53. Wang, S., Mura, C., Sawaya, M. R., Cascio, D., and Eisenberg, D. (2002) *Acta Crystallogr. D Biol. Crystallogr.* **58,** 571–578
54. Gabelli, S. B., Bianchet, M. A., Bessman, M. J., and Amzel, L. M. (2001) *Nat. Struct. Biol.* **8,** 467–472
55. Bailey, S., Sedelnikova, S. E., Blackburn, G. M., Abdelghany, H. M., Baker, P. J., McLennan, A. G., and Rafferty, J. B. (2002) *Structure* **10,** 589–600
56. Kang, L. W., Gabelli, S. B., Bianchet, M. A., Xu, W. L., Bessman, M. J., and Amzel, L. M. (2003) *J. Bacteriol.* **185,** 4110–4118
57. Kang, L. W., Gabelli, S. B., Cunningham, J. E., O'Handley, S. F., and Amzel, L. M. (2003) *Structure* **11,** 1015–1023
58. Shen, B. W., Perraud, A. L., Scharenberg, A., and Stoddard, B. L. (2003) *J. Mol. Biol.* **332,** 385–398
59. Gabelli, S. B., Bianchet, M. A., Azurmendi, H. F., Xia, Z., Sarawat, V., Mildvan, A. S., and Amzel, L. M. (2004) *Structure* **12,** 927–935
60. Ranatunga, W., Hill, E. E., Mooster, J. L., Holbrook, E. L., Schulze-Gahmen, U., Xu, W., Bessman, M. J., Brenner, S. E., and Holbrook, S. R. (2004) *J. Mol. Biol.* **339,** 103–116
61. Yoshiba, S., Ooga, T., Nakagawa, N., Shibata, T., Inoue, Y., Yokoyama, S., Kuramitsu, S., and Masui, R. (2004) *J. Biol. Chem.* **279,** 37163–37174
62. She, M., Decker, C. J., Chen, N., Tumati, S., Parker, R., and Song, H. (2006) *Nat. Struct. Mol. Biol.* **13,** 63–70
63. Zha, M., Zhong, C., Peng, Y., Hu, H., and Ding, J. (2006) *J. Mol. Biol.* **364,** 1021–1033
64. Scarsdale, J. N., Peculis, B. A., and Wright, H. T. (2006) *Structure* **14,** 331–343
65. Wakamatsu, T., Nakagawa, N., Kuramitsu, S., and Masui, R. (2008) *J. Bacteriol.* **190,** 1108–1117
66. Gabelli, S. B., Bianchet, M. A., Ohnishi, Y., Ichikawa, Y., Bessman, M. J., and Amzel, L. M. (2002) *Biochemistry* **41,** 9279–9285
67. Zha, M., Guo, Q., Zhang, Y., Yu, B., Ou, Y., Zhong, C., and Ding, J. (2008) *J. Mol. Biol.* **379,** 568–578
68. Frick, D. N., Weber, D. J., Gillespie, J. R., Bessman, M. J., and Mildvan, A. S. (1994) *J. Biol. Chem.* **269,** 1794–1803
69. Shimokawa, H., Fujii, Y., Furuichi, M., Sekiguchi, M., and Nakabeppu, Y. (2000) *Nucleic Acids Res.* **28,** 3240–3249
70. Harris, T. K., Wu, G., Massiah, M. A., and Mildvan, A. S. (2000) *Biochemistry* **39,** 1655–1674
71. Mildvan, A. S., Xia, Z., Azurmendi, H. F., Saraswat, V., Legler, P. M., Massiah, M. A., Gabelli, S. B., Bianchet, M. A., Kang, L. W., and Amzel, L. M. (2005) *Arch. Biochem. Biophys.* **433,** 129–143
72. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680

Mutation Research 718 (2011) 10–17

Contents lists available at ScienceDirect

Mutation Research/Genetic Toxicology and Environmental Mutagenesis

journal homepage: www.elsevier.com/locate/gentox
Community address: www.elsevier.com/locate/mutres

Phenylalanine 171 is a molecular brake for translesion synthesis across benzo[*a*]pyrene-guanine adducts by human DNA polymerase kappa

Akira Sassa[a,b], Naoko Niimi[a], Hirofumi Fujimoto[c], Atsushi Katafuchi[a], Petr Grúz[a], Manabu Yasui[a], Ramesh C. Gupta[d], Francis Johnson[d], Toshihiro Ohta[b], Takehiko Nohmi[a,*]

[a] *Division of Genetics and Mutagenesis, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan*
[b] *School of Life Sciences, Tokyo University of Pharmacy and Life Sciences, Hachiouji-shi, Tokyo 192-0392, Japan*
[c] *Division of Radiological Protection and Biology, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan*
[d] *Department of Pharmacological Sciences, Stony Brook University, Stony Brook, New York 11794-3400, NY, United States*

ARTICLE INFO

Article history:
Received 30 September 2010
Received in revised form 5 November 2010
Accepted 5 November 2010
Available online 13 November 2010

Keywords:
Translesion DNA synthesis
DNA polymerase κ
Benzo[*a*]pyrene
diolepoxide-N^2-deoxyguanine

ABSTRACT

Human cells possess multiple specialized DNA polymerases (Pols) that bypass a variety of DNA lesions which otherwise would block chromosome replication. Human polymerase kappa (Pol κ) bypasses benzo[*a*]pyrene diolepoxide-N^2-deoxyguanine (BPDE-N^2-dG) DNA adducts in an almost error-free manner. To better understand the relationship between the structural features in the active site and lesion bypass by Pol κ, we mutated codons corresponding to amino acids appearing close to the adducts in the active site, and compared bypass efficiencies. Remarkably, the substitution of alanine for phenylalanine 171 (F171), an amino acid conserved between Pol κ and its bacterial counterpart *Escherichia coli* DinB, enhanced the efficiencies of dCMP incorporation opposite (−)- and (+)-trans-anti-BPDE-N^2-dG 18-fold. This substitution affected neither the fidelity of TLS nor the efficiency of dCMP incorporation opposite normal guanine. This amino acid change also enhanced the binding affinity of Pol κ to template/primer DNA containing (−)-trans-anti-BPDE-N^2-dG. These results suggest that F171 functions as a molecular brake for TLS across BPDE-N^2-dG by Pol κ and that the F171A derivative of Pol κ bypasses these DNA lesions more actively than does the wild-type enzyme.

1. Introduction

Human genomic DNA is continually subject to damage caused by exogenous and endogenous genotoxic agents. The progress of the replicative DNA polymerases (Pols), such as Pol α, Pol δ and Pol ε, can be blocked by DNA lesions, possibly leading to cell death. However, these blocks can be circumvented by translesion synthesis (TLS) catalyzed by specialized Pols, or by the recombination of damaged template DNA [1,2]. Human cells possess multiple specialized Pols, including Pol η, Pol ι, Pol κ, Rev1, and Pol ζ; Pol ζ is a member of the B family and the others are members of the Y family [3]. These Pols contribute to cell survival by inserting dNMPs opposite the lesions in template DNA in an error-free or error-prone manner, thus promoting continuous chromosome replication. An example of error-free TLS is the Pol η-dependent insertion of two dAMPs opposite a cis-syn thymine DNA dimer, protecting human skin cells from ultraviolet light-induced DNA damage. Pol η deficiency leads to the genetic disease Xeroderma pigmentosum variant, resulting in increased susceptibility to ultraviolet light-induced skin cancer [4,5]. However, Pol η is suggested to be involved in error-prone TLS via its incorporation of incorrect dNMPs opposite certain DNA lesions, for example, N^2-deoxyguanosine (dG) adducts derived from 2-amino-3-methylimidazo[4,5-f]quinoline [6]. Such erroneous TLS may result in mutagenesis, carcinogenesis, and perhaps aging [7–9]. Thus, TLS can be a double-edged sword, as the mechanisms contributing to genomic integrity in the face of genotoxic agents can themselves result in mutation.

Benzo[*a*]pyrene (BP) is a ubiquitous environmental pollutant that is present in tobacco smoke and is released to the air as a product of fossil fuel combustion [10]. Upon incorporation into cells (mainly through the lungs), BP is metabolised to various reactive intermediates, the most mutagenic and tumorigenic of which is benzo[*a*]pyrene-7,8-dihydrodiol-9,10-epoxide (BPDE) [10,11]. Pol

Abbreviations: Pol, DNA polymerase; TLS, translesion DNA synthesis; dG, N^2-deoxyguanosine; BP, benzo[*a*]pyrene; BPDE, benzo[*a*]pyrene-7,8-dihydrodiol-9,10-epoxide; (+)-BPDE-N^2-dG, (+)-trans-anti-BPDE-N^2-dG; (−)-BPDE-N^2-dG, (−)-trans-anti-BPDE-N^2-dG; F171, phenylalanine 171; R175, arginine 175; L197, leucine 197; F171A (R175AL197A, D198A or E199A), a variant of hPol κ having alanine instead of F171 (R175, L197, D198 or E199); PAGE, polyacrylamide gel electrophoresis; 18C (AT or G), Alexa546-labelled 18-mer standard DNA containing dC (A, T or G) at the 13th position; 17Δ[1], 17-mer DNA lacking one base at the 13th position; 16Δ[2], 16-mer DNA lacking two bases at positions 13 and 14; K_D, the equilibrium dissociation constant; D198, aspartic acid 198; E199, glutamic acid 199.

* Corresponding author. Tel.: +81 3 3700 9872; fax: +81 3 3700 2348.
E-mail address: nohmi@nihs.go.jp (T. Nohmi).

1383-5718/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.mrgentox.2010.11.002